Plane Extraction using Point Cloud Data for Service Robot

Hiroyuki Masuta, Tatsuo Motoyoshi, Ken'ichi Koyanagi Kei Sawai and Toru Oshima

Toyama Prefectural University
Dept. of Intelligent Systems Design Engineering
Imizu, Toyama, JAPAN
masuta@pu-toyama.ac.jp

Abstract-- This paper describes an plane extraction method using point cloud data to perceive an unknown object for a service robot. Recently, depth sensors are used to perceive 3D space for a robot. A depth sensor have been used to recognize unknown environment, such as surface reconstruction, model fitting and so on. Point Cloud Library is typical open source library to deal with 3D point cloud data. However, robot perception for grasping have limitations with high computational costs and low-accuracy for perceiving small objects. Therefore, we propose the PSO-based plane detection method with RG to reconstruct an object from a combination of detected planes. To verify accuracy and computational cost for the plane detection of unknown object, we show that the proposed method has higher accuracy and less computational cost for the proposed method.

Keywords- service robot; plane extraction; depth sensor; robot vision; PSO; region growing;

I. INTRODUCTION

Recently, various types of intelligent robots have been developed for the next human society. An intelligent robot is expected to work in a familiar environment such as homes and commercial and public facilities [1]. In addition, amusement robots, robot partners and service robots will be expected in order to keep a comfortable life. These intelligent robots are expected to work in environments such as a home, a commercial and a public facility. An intelligent robot should remain in action in order to fulfill specific tasks in real time, even in an unknown environment. A robot perception to work in unknown environment is the important element.

Now, we are developing an intelligent service robot system for clearing dishes on a table. This service robot is equipped with various types of sensors for perceiving an environment. We have developed a service robot system for clearing tasks through human interaction [2]. However, a service robot system can't operate appropriately in the unexpected situation that has unknown objects. Despite this, a service robot should perceive an environmental situation and take a suitable action in real time.

A depth sensor are often used for perceiving an environment. A depth sensor like Kinect can measure a significant amount of distance data which is called point cloud data. To perceive an unknown object using point cloud data, F. Jurie et al. had proposed the real time 3D template matching to detect objects in 3D space [3]. Taylor et al. proposed a visual servoing based on object recognition applying 3D template matching using a generalized box model [4]. However, these methods need predefined knowledge which is given by the operator, or achieved by oneself. Moreover, a 3D template matching needs high computational costs because 3D space information is composed of huge amount of data. Therefore, 3D template matching is difficult to apply for robot perception, because a service robot should take action immediately in the facing situation. On the other hand, the sample consensus and segmentation methods are popular methods to detect an unknown object from point cloud data. Random sample consensus (RANSAC) and Randomized Hough Transform (RHT) are a popular segmentation method [5, 6]. For Example, K. Yamazaki et al. proposed a grasp planning method for a manipulator by estimating an object size, posture and position by using a shape model [7]. However, this method is required predefined knowledge, such as 3D CAD data, training data and so on. An appropriate predefined knowledge is difficult to provide beforehand when a service robot works in unknown environment.

Furthermore, there are some problem to apply a robot perception. For example, a service robot should make a decision in several seconds, so the acquired point cloud data of an object is imperfect data which is measured from one direction. And, it is difficult to reconstruct a accurate shape of object, because a point cloud data on an object for grasping which is measured from a robotic eye is relatively small.

Therefore, we propose a plane extraction based perceptual system for an unknown object detection from point cloud data without predefined knowledge. To detect an unknown object rapidly, we have proposed a simplified plane detection method [8]. However, a density of point cloud data on an object is not constant because the far side from a service robot is rough. Therefore, it is caused for false recognition when a service robot decide the grasping action. In this paper, we propose a plane model extraction by using a simplified plane detection method.

This paper is organized as follows. Section 2 explains

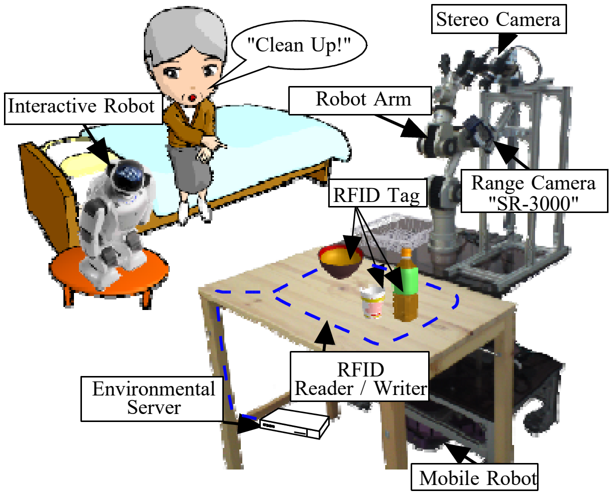


Fig.1 The service robot system overview

our intelligent service robot system and assumed measurement space. Section 3 explains our plane extraction method based on a simplified plane detection. Section 4 shows the experimental results. Section 5 concludes this paper.

II. THE ROBOT SYSTEM FOR CLEARING TABLE

Fig. 1 shows an overview of the intelligent service robot system for clearing a table. The intelligent robot system consists of a service robot, an interactive robot, and an intelligent space server [2]. The interactive robot recognizes human orders from hand gestures and spoken commands using a stereo vision system and a voice recognition system. The intelligent space server manages the information on table objects, such as dish type, size, and color, using a Radio-Frequency Identification (RFID) system. The service robot has a robot arm, and the mobile robot collects a dish based on vision information. However, an RFID reader cannot read RFID tags, and the vision cannot recognize unregistered objects in unexpected situations that have unknown objects. An intelligent robot should detect workable objects from unregistered objects.

We installed SR-3000 as a depth sensor to measure the environment. SR-3000 is a Time-Of-Flight range image sensor made by MESA imaging AG that can measure 3D distance and 2D luminance. The range of this sensor is a maximum of 7.5 [m] [9], and the output signal is a quarter common intermediate format (QCIF 176 * 144 pixels). Therefore, this sensor can measure a 25,344-directional distance information simultaneously. SR-3000 is located at the height of 60 [cm] from the base of the robot arm to fit certain objects such as tables into the sensor range.

Therefore, an intelligent service robot can measure a significant amount of data simultaneously. An intelligent service robot have to reduce the computational cost for taking real time action. Fig. 2 shows a snapshot of the

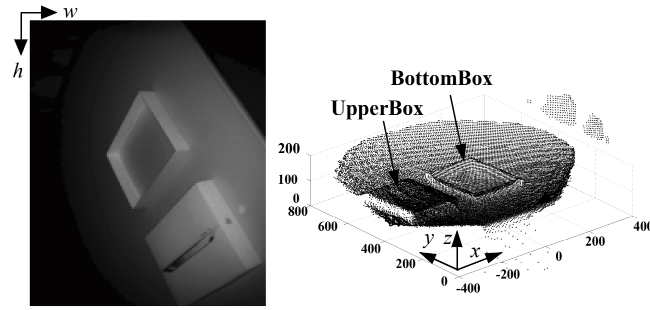


Fig.2 The snapshot of SR-3000 image

distance and amplitude image of the product box. We define the measuring point of the image pixel as the *w*-*h* axis of the coordinate. Moreover, the measuring point of the real world is defined as the *x*-*y*-*z* axis of the robot coordinate, where the origin is the base of the robot arm.

The upper box and bottom box are recognized from point cloud data by observation. However, the distance data of the box surface is not completely. The left side of upper box is not measured by occlusion and the boundary of the surface is smooth by smoothing filter. Furthermore, the edge of measured area is actually curved for the lens aberration. Therefore, above all conditions should be considered to perceive the unexpected situation. To detect a grasping object from this sensing data, we focus on plane detection method, because we consider a combination of some planes to be an object.

III. PLANE EXTRACTION FROM POINT CLOUD DATA

A. Simplified Plane Detection Method with PSO

This section explains a plane detection method to extract a point cloud group of a plane on an unknown object. Fig.3 shows the processing flow of the proposed method from measurement to decision making. Firstly, Particle Swarm Optimization (PSO) is applied to select appropriate reference point for the Simplified Plane Detection (SPD). This subsection explains the plane detection method for the unknown object detection using a point cloud data. A typical plane detection methods are Hough transform (HT) and random sample consensus (RANSAC). HT can find imperfect instances of objects within a certain category of shapes through a voting process. However, this voting process requires a high computational cost. The Randomized Hough Transform (RHT) is proposed to reduce the computational cost more than HT. RHT reduces the voting process by selecting a candidate plane randomly. But, RHT is not sufficient to satisfy online plane detection for robot perception. RANSAC is another well-known robust model fitting method [5]. RANSAC is an iterative method to estimate likely plane parameters from a set of observed data that contains to inliers. However, using this method, it is difficult to detect the planes of small objects with certainty.

We have proposed a simplified plane detection method

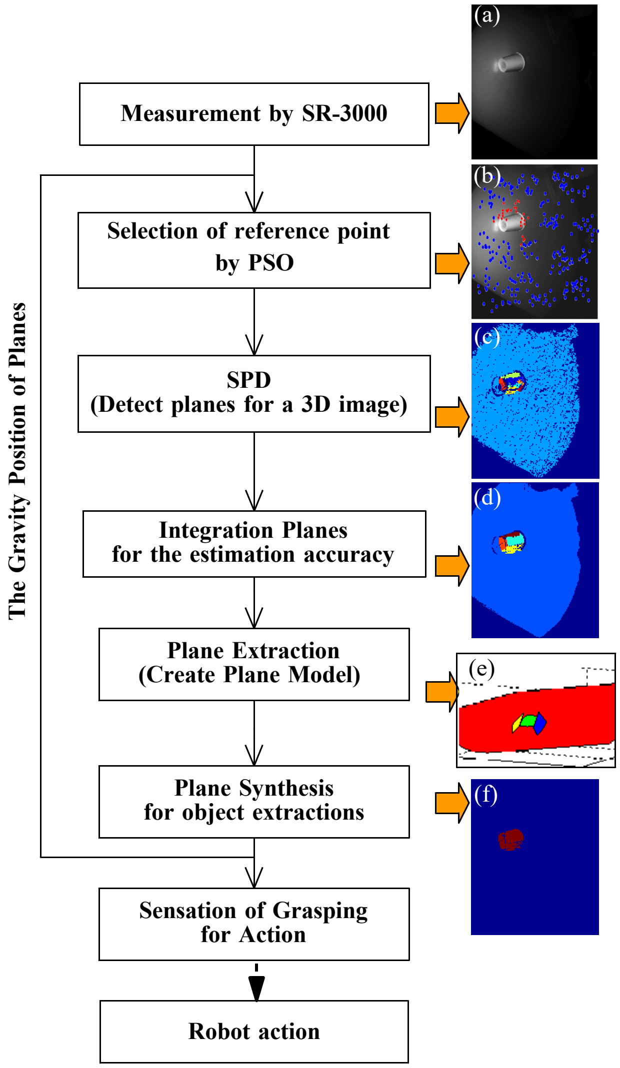


Fig. 3 The processing flow for deciding robot action

(SPD) that excludes the voting process and iterative process in order to reduce computational costs. Fig. 3 shows a coordinate system in the measurement space if a depth sensor measures a cubic object. First, we evaluate the probability of a plane from vectors around an arbitrary reference point $P = (x_P, y_P, z_P)$ as follows:

$$V = \left| \mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3 \right| \tag{1}$$

Volume V is calculated from vectors $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ around an arbitrary reference point. These vectors are on the same plane Π when V is nearly equal to zero. We calculate the normal vector $\mathbf{n} = (n_x, n_y, n_z)$ of this plane as follows:

$$\mathbf{n} = \mathbf{v}_1 \times \mathbf{v}_2 \approx \mathbf{v}_2 \times \mathbf{v}_3 \approx \mathbf{v}_3 \times \mathbf{v}_1 \tag{2}$$

where θ and φ denote the two angles (azimuth and elevation) associated with the spherical representation of the plane's unit length normal vector. This plane's posture

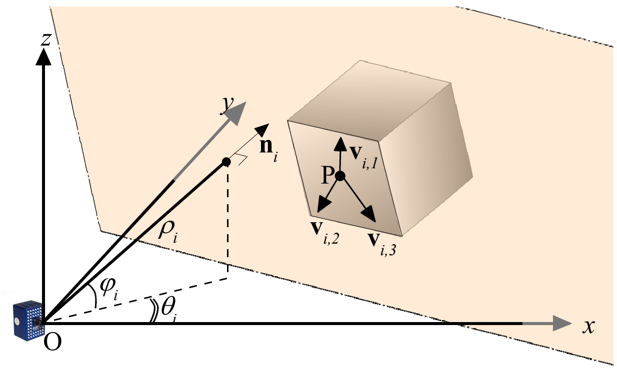


Fig. 4 Relations of parameters for plane detection

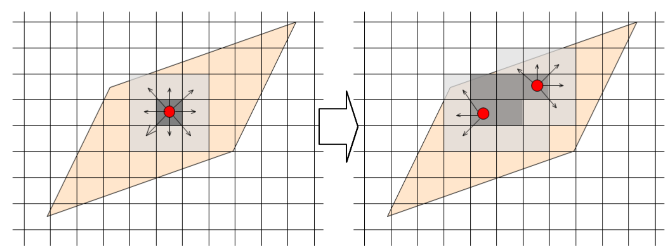


Fig. 5 Region growing for plane detection

$[\rho, \theta, \varphi]$ is calculated based on equation (3).

$$\begin{cases} \varphi = \sin^{-1}(n_z) \\ \theta = \cos^{-1}\left(\dfrac{n_x}{\cos\varphi}\right) \\ \rho = x_P \cos\theta\cos\varphi + y_P \sin\theta\cos\varphi + z_P \sin\varphi \end{cases} \tag{3}$$

Equation (3) is often used in the Hough transform. Next, the region growing (RG) method is applied to obtain the measuring points that belong in the plane Π. Fig. 4 shows an image of the RG. We ascribe the arbitrary reference point to a seed for RG [10]. First, the surrounding points of the arbitrary point that belong in the plane Π are searched as satisfying equation (4).

$$\mathbf{n}_{ref} \cdot \mathbf{n}_{ref+k} \cong 1.0 \quad \wedge \quad \rho_{ref} \cong \rho_{ref+k} \tag{4}$$

where $\mathbf{n}_{ref}$ is the normal vector of the arbitrary point, $\mathbf{n}_{ref+k}$ is the normal vector of the k-th searching point, ρ_{ref} is equation (3) of the arbitrary reference point, and ρ_{ref+k} corresponds to k-th searching point. If a new point that belongs in the plane Π is found, the points that surround the new point are searched repeatedly. When no point that belongs in the plane Π cannot be found, RG for plane Π is finished. The gravity point of a detected plain is calculated from a set of points, and the location of the plane is calculated from the distance information. Other planes are detected through this repetitive procedure. Therefore, SPD can detect a homogeneous plane without the voting process. In our method, the important point for reducing the computational cost and for obtaining small planes is the selection of arbitrary reference points. We apply the particle swarm optimization (PSO) to set appropriate seeds.

B. PSO for selecting appropriate reference point

PSO was developed by Kennedy and Eberhart in 1995, and it was inspired from the social behavior of a flock of migrating birds attempting to reach an unknown destination [11]. We expected that PSO balances accurate plane detection with the reduction of computational costs. The process of PSO is initialized with a group of random particles. Particle i is represented as a point in the N-dimensional space in the search space S, where N is the number of variables. Throughout the process, each particle i monitors three values: its current position ($\mathbf{x}_i$), the best position it has reached in previous cycles ($\mathbf{pbest}_i$), and its flying velocity ($\mathbf{v}_i$). The process is represented as follows:

$$\begin{cases} \mathbf{x}_{in} \equiv (x_{i1}, x_{i2}, \cdots, x_{iN}) \\ \mathbf{v}_{in} \equiv (v_{i1}, v_{i2}, \cdots, v_{iN}) \\ \mathbf{pbest}_{in} \equiv (pbest_{i1}, pbest_{i2}, \cdots, pbest_{iN}) \end{cases} \tag{5}$$

After initialization, plane detection is performed for particles as seeds for RG. The evaluation of each particle is calculated following the fitness function:

$$f_k(\mathbf{x}_{in}(t)) = \frac{1}{2\pi\sigma^2} \exp\left\{ -\frac{(h - hg_k)^2 + (w - wg_k)^2}{2\sigma^2} \right\} \quad (f_k \in \Pi_k) \tag{6}$$

where h_g and w_g are the image gravity position that was a detected plane in a previous step. σ is standard deviation. The particle that has a large fitness value is acceptable. At each time interval (*cycle*), the position of the best particle (**lbest**) is calculated for K particles as the best fitness of all particles, where K is the number of particles. Velocity $\mathbf{v}_i$ is updated as follows:

$$\begin{aligned} \mathbf{v}_{in}(t+\Delta t) = & w * \mathbf{v}_{in}(t) + c_1 * rand_1 * (\mathbf{pbest}_m - \mathbf{x}_{in}(t)) \\ & + c_2 * rand_2 * (\mathbf{lbest}_n - \mathbf{x}_{in}(t)) \end{aligned} \tag{7}$$

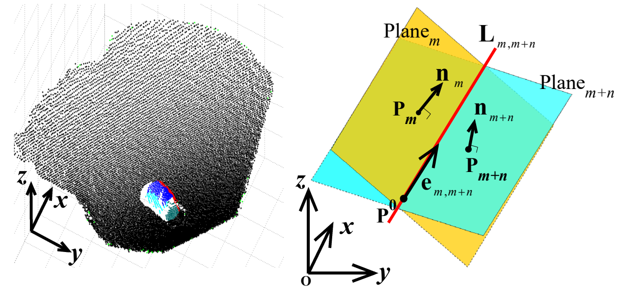
Each particle position is updated using the new velocity $\mathbf{v}_i$:

$$\mathbf{x}_{in}(t+\Delta t) = \mathbf{x}_{in}(t) + \mathbf{v}_{in}(t+\Delta t) \tag{8}$$

For these processes, the particles gather around the detected plane. However, the search area is limited to the specific plane. Therefore, particles that have low fitness values are mutated when the number of detected planes is larger than ε. The particles are to be the reference point for SPD.

C. The Integration Planes for The Estimation Accuracy

Fig. 3(a), (b) and (c) shows a snapshot of SR-3000 image, distribution of particles and a result of plane detection, respectively. However, it is difficult to detect a steady plane because of sensing noise. We apply the estimation accuracy parameter to consider the changing time series as shown by the following equation.



(a)Result of plane detection (b)Definition of plane integration

Fig. 6 Plane Integration for detecting object

$$\begin{cases} W_{w,h}(t) = 0.5 & \text{if (detect a new plane)} \\ W_{w,h}(t) = \gamma \cdot W_{w,h}(t-1) + k & \text{if (detect a plane the same as previous plane)} \\ W_{w,h}(t) = \gamma \cdot W_{w,h}(t-1) - r & \text{if (detect a plane the different previous plane)} \\ W_{w,h}(t) = \gamma \cdot W_{w,h}(t-1) & \text{if (detect no plane)} \end{cases} \tag{9}$$

where $W_{w,h}$ is the estimation accuracy parameter for each pixel, γ is decrease rate, and k and r are constant values. If a detected plane is new, $W_{w,h}$ is initialized to 0.5. Then, $W_{w,h}$ is increased if the detected plane is the same as the previously detected plane. $W_{w,h}$ is decreased if a detected plane is different from the previously detected plane. $W_{w,h}$ is slowly decreased with time if there is no plane. Hence, the estimation accuracy parameter is increased when the same plane is detected continuously. The steady plane is switched to the detected plane or eliminated when the $W_{w,h}$ is smaller than the threshold. As a result, a stable plane detection result is obtained, similar to Fig. 5(d).

D. Plane Extraction based on Plane Detection

Above plane detection method detected point cloud data on the plane. However, the number of point cloud data is different according to distance from a depth sensor to an object. Hence, the density of a point cloud data is also different even if same plane size. We apply a convex-hull algorithms to extract an edge of plane [12]. A posture of a plane that calculated by SPD is not accurate, because it is considered only neighbor vectors around the arbitrary reference point. Here, we applied RANSAC algorithm to correct posture of a plane from detected point cloud data. The processing is as follows;

1. select a random subset(4 points) of the detected point cloud data of a plane. Subset calls the hypothetical inliers.
2. A plane model is calculated from subset using equation (1) to (3).
3. All other data are tested against the hypothetical inliers.
4. Through repetition above process, the estimated plane model is reasonably good if sufficiently many points are clustered in the inlier.

As a result, 3D planes are extracted like Fig.3(e). The computational costs could be reduced by narrow down the candidate point cloud data on a plane.

E. Plane synthesis for object extraction

There are many planes on the object position in the case of a round object and multi-object situation like Fig.6(a). Therefore, some planes are integrated to one plane set as an object. We focus on geometric invariance of each plane for object extraction.

Fig.6(b) is shown the definition of the two different plane integration. First, the equation of the plane is calculated for each detected plane. The unit vector of a line of intersection between the m-th plane and the $(m+n)$-th plane is calculated as follows.

$$\mathbf{e}_{m,m+n} = \mathbf{n}_m \times \mathbf{n}_{m+n} \tag{10}$$

The line of intersection vector is shown by equation (11).

$$\mathbf{L}_{m,m+n} = \mathbf{P}^0 + t\mathbf{e}_{m,m+n} \tag{11}$$

where t is parameterization, $\mathbf{P}^0$ is the arbitrary point on the line of intersection. Next, the line of intersection is evaluated, where it passes on the m-th plane and $(m+n)$-th plane using the following equation.

$$\begin{aligned} r_m &= \sum_{j=1}^{R} \left| (\mathbf{P}^0 + t_{m,j}\mathbf{e}_{m,m+n}) - \mathbf{P}_{m,j} \right| \\ t_{m,j} &= \frac{(\mathbf{P}_{m,j} - \mathbf{P}^0)\mathbf{e}_{m,m+n}}{\left|\mathbf{e}_{m,m+n}\right|^2} \end{aligned} \tag{12}$$

where $P_{m,j}$ is a whole series of point of the m-th plane. The number of points belonging to the m-th plane is R. Finally, two planes are integrated, if r_m *and* r_{m+n} are smaller than the threshold. An object is observed as one plane set by repetition.

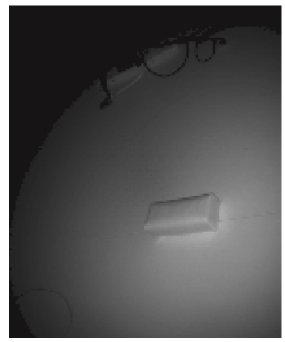
IV. EXPERIMENT

Experiments were conducted by the proposed method under two conditions. The robot position is determined. The object information is not provided to the robot. This means the environmental situation is unknown to the robot. The number of total particles and local best particles are 200 and 20, respectively. The specifications of the computer used are as follows. CPU: Core i7 5600U with 2.60GHz (4 CPUs); Memory (RAM): 8 GB; Software: MATLAB/Simulink.

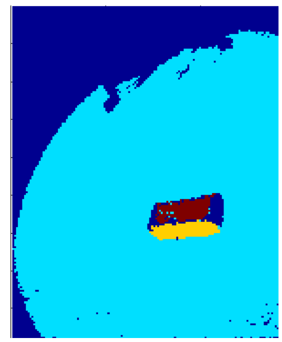
Fig. 7 shows experimental result on a single object condition. Fig.7(a) shows experimental condition that there is a confectionary boxes on the table. Fig.7(b) shows a snapshot of SR-3000 image. Fig.7(c) shows a result of a plane detection. A color is corresponding to an identical detecting plane. There are 3 planes, and top plane and side plane on the box are detected separately. Fig.7(d) shows a result of all point cloud data with plane detection result. The density of point cloud data of far side of a detected plane from SR-3000 is lower. Fig.7(e) shows a result of a plane extraction. Each detected plane is reconstructed to a plane model.

The computational time of the 50 steps average on the proposed method is 0.931 [s], which the plane detection is 0.045 [s] and the plane extraction is 0.886 [s]. Therefore, it is expected to reduce computational cost, if the faster convex-hull algorithm is applied.

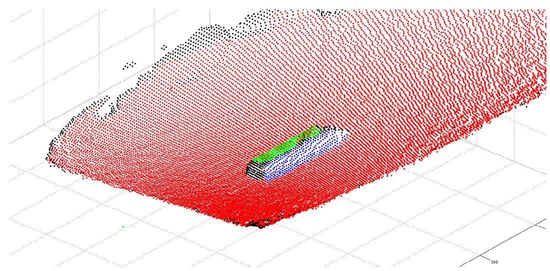
(a) Experimental condition



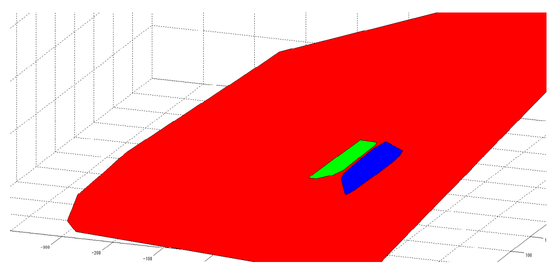
(b) Snapshot of SR-3000 image



(c) The result of plane detection



(d)The result of plane detection on point cloud data



(e) The result of plane extraction

Fig. 7 Experimental Result on a single object

Fig.8 shows experimental result on a multiple objects condition. Fig.8(a) shows experimental condition that there are 6 boxes on a table. Fig.8(b) shows a snapshot of SR-3000 image. Fig.8(c) shows a result of a plane detection. Planes on boxes and table surface are detected separately. Fig.8(d) shows a result of all point cloud data with plane detection result. There is deficiency of a point data in some detected planes. This reasons is affected by a sensing noise. Fig.8(e) shows a result of a plane extraction. In this plane model is complement a deficiency of a point data.

Furthermore, the proposed method has a merit that can consider the continuity even if a part of plane is hidden by other object.

V. CONCLUSION

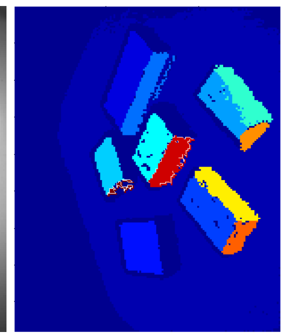
This paper describes an plane extraction method using point cloud data to perceive an unknown object for a service robot. A depth sensor have been used to recognize unknown environment, such as surface reconstruction, model fitting and so on. However, a service robot system can't operate appropriately in the unexpected situation that has unknown objects. Despite this, a service robot should perceive an environmental situation and take a suitable action in real time. To perceive an unknown object using point cloud data without predefined knowledge, we have proposed the PSO-based plane detection method with RG, and the plane extraction method applied with convex-hull algorithm.

To verify accuracy and computational cost for the plane extraction of unknown object, we show that the proposed method has higher accuracy and less computational cost for the proposed method. In multiple objects situation, we show that the plane models are reconstructed the environment from a complicated point cloud data. Furthermore, the proposed method has a merit that can consider the continuity even if a part of plane is hidden by other object.

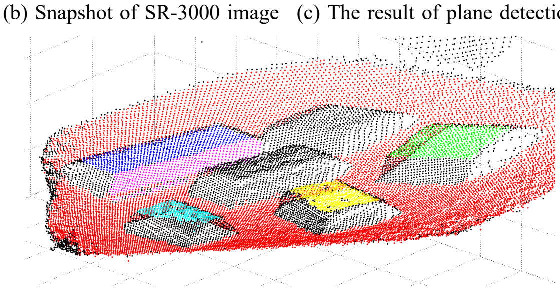
For future work, we will apply the high speed calculation for convex-hull algorithm. And, we will verify the robot action through all the process in Fig.3.

(a) Experimental condition

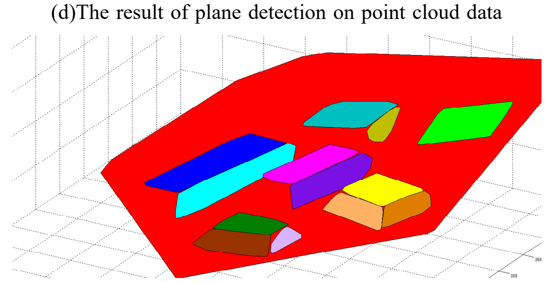
(b) Snapshot of SR-3000 image



(c) The result of plane detection



(d)The result of plane detection on point cloud data



(e) The result of plane extraction

Fig. 8 Experimental Result on a multiple objects

REFERENCES

[1] M.Beetz, U.Klank, I.Kresse, A.Maldonado, L.Mosenlechner, D.Pangercic, T. Ruhr, M.Tenorth: "Robotic roommates making pancakes" 11th IEEE-RAS International Conference on Humanoid Robots, pp529-536, 2011.

[2] H.Masuta, E.Hiwada and N.Kubota, "Control Architecture for Human Friendly Robots Based on Interacting with Human", 4th International Conference on the 2011 International Conference on Intelligent Robotics and Applications, Part II, LNAI 7102, pp. 210-219, 2011.

[3] F.Jurie, M.Dhome, "Real time 3D template matching", Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol.1, pp.I-791-796, 2001.

[4] G.Taylor, L.Lindsay, "Grasping Unknown Objects with a Humanoid Robot", Proceedings of 2002 Australasian Conference on Robotics and Automation, pp.191-196, 2002.

[5] M.A.Fischler and R.C.Bolles: "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography", Communications of the ACM, Vol 24, Issue 6, pp 381-395, 1981.

[6] L.Xu and E.Oja: "Randomized Hough transform (RHT): basic mechanisms, algorithms, and computational complexities", CVGIP Image Understanding, Vol.57, No.2, pp131-154, 1993.

[7] K.Yamazaki, M.Tomono, T.Tsubouchi, S.Yuta, "A grasp planning for picking up an unknown object for a mobile manipulator," Proceedings 2006 IEEE International Conference on Robotics and Automation, pp.2143-2149, 2006.

[8] H.Masuata, H.Lim, "Direct Perception for Intelligent Robot to Grasp an Unknown Object", Proceedings of the 3rd International Workshop on Advanced Computational Intelligence and Intelligent Informatics, SS1-3, 2013.

[9] T. Oggier, M. Lehmann, R. Kaufmannn, M. Schweizer, M. Richter, P. Metzler, G. Lang, F. Lustenberger, N. Blanc, "An all-solid-state optical range camera for 3D-real-time imaging with sub-centimeter depth-resolution (SwissRanger)", Proceedings of SPIE Vol. 5249, 2003, pp.634-545.

[10] R.Adams, L.Bischof: "Seeded Region Growing", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.16, pp. 641-647, 1994.

[11] J. Kennedy, R. C. Eberhart: “Particle swarm optimization”, In Proceedings of the 1995 IEEE International Conference on Neural Networks”, Vol. 4, pp. 1942-1948, 1995.

[12] R. L. Graham, "An Efficient Algorithm for Determining the Convex Hull of a Finite Planar Set", Information Processing Letters 1, 132-133, 1972.