2015 IEEE Symposium Series on Computational Intelligence

Invariant Perception for Grasping an Unknown Object using 3D Depth Sensor

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

Hun-ok Lim
Kanagawa University
Dept. of Mechanical Engineering
Yokohama, Kanagawa, JAPAN

Tatsuo Motoyoshi, Ken'ichi Koyanagi and Toru Oshima
Toyama Prefectural University
Dept. of Intelligent Systems Design Engineering
Imizu, Toyama, JAPAN

Abstract-- This paper describes robot perception to grasp an unknown object intuitively. Generally, a robot should recognize object's property such as size, posture and position to make a decision of a grasping task. However, an accurate object information is difficult to perceive quickly from limited measured information, when a robot perceive an unknown object. So, a robot is difficult to decide a suitable action. Therefore, we propose a robot perception system to decide a suitable action intuitively from 3D point cloud data from limited information. We propose robot perception system which is composed of an online processable object extraction method and an invariant detection method for making a decision of a grasping action. The object extraction method can extract a part of point cloud corresponding to an object from 3D point cloud data. The invariant detection method for a grasping action is explained by inertia tensor and fuzzy inference. The invariant for a grasping action affords the possibility of action to a robot directly without inference from object's property such as size, posture and shape. As experimental results, we show that the object extraction reduces a computational cost drastically, but also the accuracy is better, further the robot can detect a relevant information of a grasping behavior directly.

I. INTRODUCTION

Recently, various types of intelligent robots have been developed for the next human society [1]. For example, security robots, rehabilitation robots, and medical robots will be useful. In addition, amusement robots, partner robots, and service robots will be expected in order to live a comfortable life. These intelligent robots are expected to work in a familiar environment such as homes and commercial and public facilities [2]. In a real environment, an intelligent robot should remain in action in order to fulfill specific tasks, even in an unknown environment. Therefore, an intelligent robot should perceive the environment flexibly, similar to humans.

There are many previous studies that focus on the perception of unknown objects and the decision making of an action in real environments. To perceive an unknown object, F. Jurie et al. had proposed the real time 3D template matching to detect objects in 3D space [3]. The sample consensus and segmentation methods is a popular method to detect the unknown object from point cloud data. Random sample consensus (RANSAC) is a popular segmentation method [4]. By using an above object detection method, K. Yamazaki et al. proposed a grasp planning method by estimating an object size, posture and position by using a shape model [5]. Taylor et al. proposed a visual servoing based on object recognition applying 3D template matching using a generalized box model [6].

On the other hand, a robot should obtain a significant amout of data from various angles to get acurate results of 3D template matching. However, a robot which moves around to measure a data in unknown environment is not realistic because it needs much time. Furthermore, a point cloud of an object is imperfect, for example a data of behind side is not found when a robot measures from one direction. Therefore, the unknown environment has a hard problem that a robot should decide to grasp an object promptly from imperfect data.

On the other hand, human obtains a significant amount of data from sensory organ. However, they can perceive the suitable situation and take suitable action promptly. An ecological psychology insists on a perceiving-action cycle which focused on the cycle of perception and action [7]. The perceiving-acting cycle in ecological psychology insists that the perceptual system is the process of picking up or detecting invariant information in an environment [8]. Here the importance is that a perception is not inferred from sense organs, but they detect only an invariant information. The concept of perceiving-acting cycle maintains the coupling of perceptual system and action system. The perceptual system and the action system restrict each other through interaction with the environment. When the perceiving-acting cycle

978-1-4799-7560-0/15 $31.00 © 2015 IEEE
DOI 10.1109/SSCI.2015.28

IEEE computer society

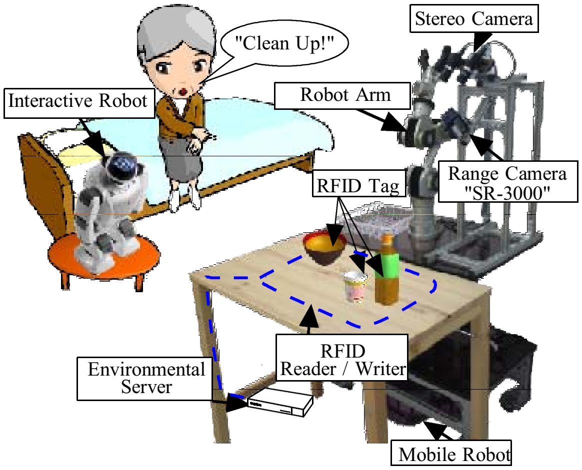

Fig.1 The service robot system overview

forms a coherent relationship with the environment, the specific perceptual information generates the specific action outputs like reactive motions. We consider the couple of a perceptual system and action system situated in the facing environment as one phase of the perceiving-acting cycles [9].

Therefore, we propose a robot perception system to make decision of a suitable action intuitively from 3D point cloud data based on perceiving-acting cycle concept. To detect an unknown object from a 3D range information promptly, we have proposed a plane detection method [10]. The plane detection method can detect small planes accurately and computational cost is less than previous methods drastically. Furthermore, we propose an invariant detection method for grasping an unknown object from 3D point cloud data of an unknown object obtained by the plane detection method.

This paper is organized as follows. Section 2 explains our intelligent robot system. Section 3 explains the plane detection method and object extraction from 3D point cloud data. Section 4 shows the invariant detection method, Section 5 shows experimental results. Finally, Section 6 concludes this paper.

II. THE ROBOT SYSTEM FOR CLEARING TABLE

A. Outline of the Service Robot System

Fig.1 shows an overview of the intelligent robot system for clearing a table. The intelligent robot system consists of a service robot, an interactive robot, and an intelligent space server [11]. In this system, the interactive robot recognizes human orders from hand gestures and spoken commands [12]. The intelligent space server manages the information on table objects, such as dish type, size, and color, using a Radio-Frequency Identification (RFID) system [13]. The service robot which has a robot arm, and a mobile robot picks up dishes based on vision information [14]. The robot arm called Katana is small and lightweight, and is similar in size to the human arm. However, object information is provided beforehand by a human operator. The robot can not pick up an object, if there are unknown objects.

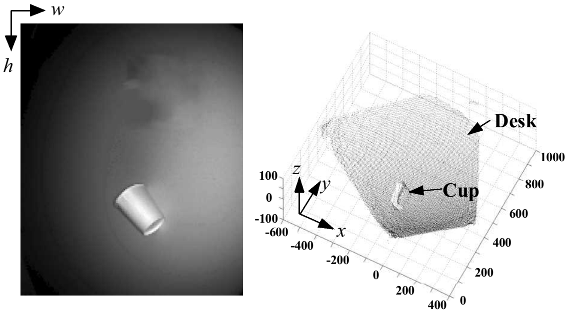

Fig.2 The snapshot of SR-3000 image

To perceive an unknown object, we installed a 3D depth sensor to the service robot. It is important to decrease a computational cost because a 3D depth sensor can measure a significant amount of data. To reduce computational cost, the researches of a region of interest (ROI) are being discussed for example characteristic pattern, change and so on [15]. In these researches, a robotic action is not considered to restrict ROI. In our research, Katana has a five degree-of-freedom (5DOF) structure and six motors, including a gripper. Therefore, Katana movement is restricted to approach from the front side. An action is restricted by the embodiment of robotic structure. Therefore, the information that is needed to perceive is also restricted by action.

B. 3D Depth Sensor for Measuring Environment

We use SR-3000 as the 3D depth sensor to measure the environment. SR-3000 is a Time-Of-Flight range image sensor made by the MESA imaging AG that can measure 3D distance and 2D luminance. The range of this sensor is a maximum of 7.5 [m] [16], 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 50 [cm] from the base of the robot arm to fit certain objects such as tables into the sensor range. The maximum frame rate of this sensor is 50 [fps], but we set a fixed frame rate of 20 [fps] because the sensing stability and real time sensing are considered.

Therefore, an intelligent robot can measure a significant volume of data simultaneously. An intelligent robot should be able to reduce the computational cost. Fig.2 shows a snapshot of the distance and amplitude image of a paper cup. 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.

A cup is recognized from a 3D distance data by observation. However, the point cloud data of the round surface is not complete. One side of the cup is not measured by occlusion, and the surface boundary is smoothed with a smoothing filter. Furthermore, the edge of the measured field is curved because of lens distortion. Therefore, all the aforementioned conditions should be considered in order to

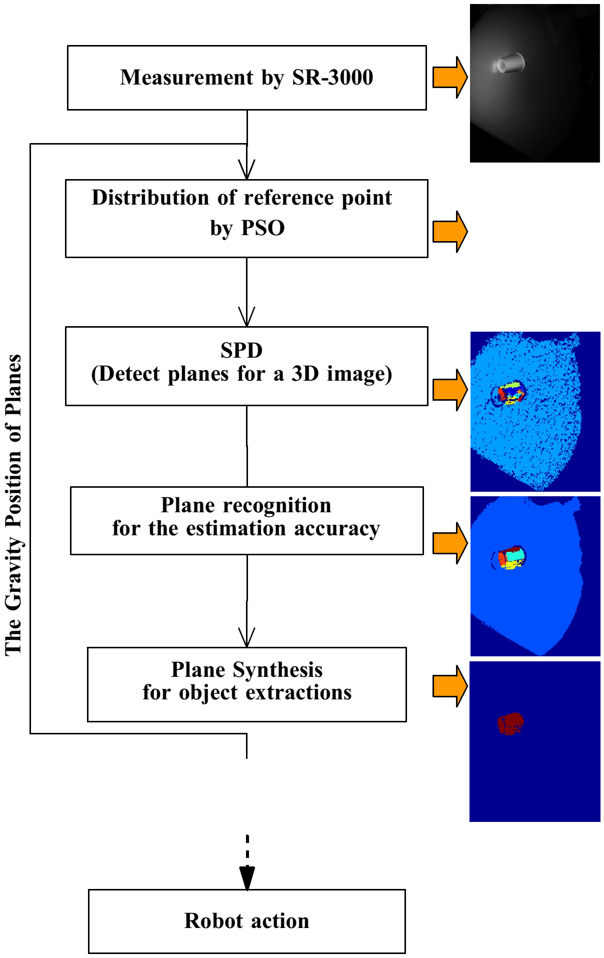

Fig.3 The processing flow of SPD-SE

perceive unexpected situations and make a decision for action. To detect an object to grasp from this sensing data, we focus on the plane detection method.

III. UNKNOWN OBJECT EXTRACTION BASED ON PLANE DETECTION IN 3D SPACE

A. The Overview of the Unknown Object Extraction Method

This section explains an unknown object extraction method for extracting a point cloud group of an unknown object. We have proposed the Simplified Plane Detection based SEgmentation of a point cloud into object's clusters (SPD-SE). Fig.3 shows the processing flow of the proposed method with SPD-SE. The SPD-SE consists of the particle swarm optimization (PSO), the simplified plane detection (SPD) with region growing (RG) and the plane integrating object extraction. The PSO for the seeded RG is the first step to detect planes which is composed of object from point cloud data. Fig.3(a), (b) and (c) shows a snapshot of SR-3000 image, distribution of particles of PSO and a result of simplified plane detection, respectively. The particles of PSO are updated to gather around detected small planes.