The Grasping of the Novel Objects given on Active Vision Systems

1. Motivation

In the era of artificial intelligence, robotics integrated technologies from different disciplines, spanning cartography to computer graphics, and machine vision algorithms, so that the application scope of robots becomes much larger. In the industrial application scenarios, robots can even achieve the grasping and classification of object, which are sometimes unknown for the robot. How to make robots reliably grasp novel object becomes an important issue. However, it is not easy to finish this task. There are collisions in the actual scenarios, by which the partial object may be hidden. In addition, it is also difficult to get the object model, which is suitable for robotic grasping. In order to model the novel object properly, we consider the use of active vision. It can support automatic grasping of objects by selecting modeling views. Given an active vision system, we would like to minimize the number of views taken, while maximizing grasping reliability. [1] How to keep balance is a challenge for us.

2. Problem Definition

In general, the main purpose is to make the robot grasp the novel objects and to evaluate the experimental results. In the grasping process, the aspects to consider are the object modeling and obtaining the safe path for grasping. Of course, in the process of modeling, what we need is not the whole object structure, but the part which can become the basis of the grasping process. Meanwhile we combine the concept of active vision. Active sensor which determines the viewpoints with a viewing strategy thus becomes critically important for achieving full automation and high efficiency.[2] Following we need active sensor planning to complete the task. That is, we need to develop intentional actions to complete the reconstruction of unknown objects. Overall, we can separate the whole task into two parts. One is the modeling of the novel object, and the other is the development of a safe and reliable grasping path based on the object model. The following describes the detailed ideas for implementation.

Firstly, what we need to accomplish is to get the grasping model of the novel objects. According to the current research, there have already been the methods for this sub-task. According to the method in [1], we can firstly model the part of the novel object with a single view, in order to obtain some candidate contact points. Then based on these candidate points, we can drive the active vision to model their neighbors until a reliable grasping model is obtained. So we should focus on the next best view exploration procedure. Of course, in this process, there may be the possibility that we cannot find a valid grasping model based on the existing information. Because what we get is only the part of the object. To solve this problem, the robotic should look at the workspace around the recovered point cloud to fulfill safety exploration. Besides the next best view procedure, the other aspect we need to consider is grasping plan. Referred to the method described in [3], two models should be learnt. One model is the contact model of the relation between a finger part and the local object shape at its contact, and the other is the hand configuration model from the example grasp. Then these two models are combined for the grasping plan. So far, we have completed the first sub-task. After obtaining the reliable grasping model, we need to find the safe reliable path for robotic grasping. We hope that the robot can complete the collision-free grasp. So we estimate the possibility of unexpected collision on the grasping path. It also means, that we select views so as to minimize the entropy of the voxels through which the robot hand is going to pass when following a given grasp trace. Based on this principle, we will choose the best path to grasp the novel object. After completing the whole task, we also need to evaluate this method. The method can be evaluated from two perspectives. One is to model the novel objects from the single view, and the other is to use our method to obtain the partial shape completion for grasping. Based on these two perspectives, the efficiency of robotic grasping will be compared. What we can compare include the average success rate per object grasping, the run time, and the average collision probability per object grasping.

3. Related Work

Nowadays, there have been studies on robots grasping of novel objects. In [4], based on the assumption that most objects have symmetries, a method for completing object shapes are proposed. Compared with that, we have proposed the concept of active vision. An active vision system is one that can manipulate the viewpoint of the camera(s) in order to investigate the environment and get better information from it.[5] Active vision makes the robot grasping process more automated, and our focus is also moved to the view selection. For the next best view selection, there have also been many studies. In [6], the detailed implementations of the Pianetarium Algorithm and the Normal Algorithm are proposed. They are both for the next best view selection. It is worth noting, that for the representation of contact model the Octree structure is used. Octrees are 8-ary trees which recursively divide a cube into eight sub cubes. The state of the current point is also included in the Octree structure, such as empty, occupied, and unseen. We will also use this structure in the proposed implementation. On the other hand, the used hardware in the experiment is robotic ARMAR-6[7]. In order to be able to adequately control the high-performance actuators of ARMAR-6,a robot-independent, real-time capable control framework was developed. It is also the fundamental of the task.

4. Schedule

	Task	Duration
1	Familiarizing myself with the platform, ARMAX-6 and understanding some basic concepts	3 Weeks
2	Sub-Task: the modeling of novel object for grasping(NBV)	4 Weeks
3	Sub-Task: the grasp planning	3 Weeks
4	Combination of step 2 and step 3	1 Weeks
5	Sub-Task: the grasp path	3 Weeks
6	Combination of step 4 and step 5	2 Weeks
7	Experimental test and evaluation analysis	6 Weeks
8	Summarizing and writing the thesis	4 Weeks
	Total	26 Weeks / 6 Months

Reference

[1] E. Arruda, J. Wyatt and M. Kopicki, "Active vision for dexterous grasping of novel objects," *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Daejeon, 2016, pp. 2881-2888.
[2] Chen, Shengyong & Li, Youfu & Ming Kwok, Ngai. (2011). Active Vision in Robotic Systems: A Survey of Recent Developments. I. J. Robotic Res.. 30. 1343-1377.
[3] Kopicki, M., Detry, R., Adjigble, M., Stolkin, R., Leonardis, A., & Wyatt, J. L. (2016). One-shot learning and generation of dexterous grasps for novel objects. *The International Journal of Robotics Research*, *35*(8), 959–976
[4] D. Schiebener, A. Schmidt, N. Vahrenkamp and T. Asfour, "Heuristic 3D object shape completion based on symmetry and scene context," *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Daejeon, 2016, pp. 74-81.
[5] https://en.wikipedia.org/wiki/Active_vision
[6] C. Connolly, "The determination of next best views," Proceedings. 1985 IEEE International Conference on Robotics and Automation, St. Louis, MO, USA, 1985, pp. 432-435.
[7] Asfour, Tamim & Kaul, Lukas & Wächter, Mirko & Ottenhaus, Simon & Weiner, Pascal & Rader, Samuel & Grimm, Raphael & Zhou, You & Grotz, Markus & Paus, Fabian & Shingarey, Dmitriy & Haubert, Hans. (2018). ARMAR-6: A Collaborative Humanoid Robot for Industrial Environments. 447-454.