Object Recognition and Pose Estimation from RGB-D Data Using Active Sensing

Udaka A. Manawadu, Shishiki Keito and Naruse Keitaro

Robot Engineering Lab, University of Aizu

Introduction and Novelty

3D object recognition and pose estimation using RGB-D cameras have become a trending topic among researchers since Point Cloud (PC) taken from RGB-D cameras contains colors, not like Light Detection and Ranging (LiDAR) sensors. Many research for object recognition and pose estimation were done from RGB-D data from only one angle [2]. Detecting objects from many angles is novel in this research. The task of automatically opening and closing a globe valve using a robot arm from different angles is considered in this research.

As in figure 1 the valve is consisting of two parts, the handle part, and the body part. from 0° to 30° was taken as the front mode, and from 30° to 90° was taken as the side mode. Intel Realsense D435 camera attached to JACO arm with six degrees of freedom was used in this research as the hardware.

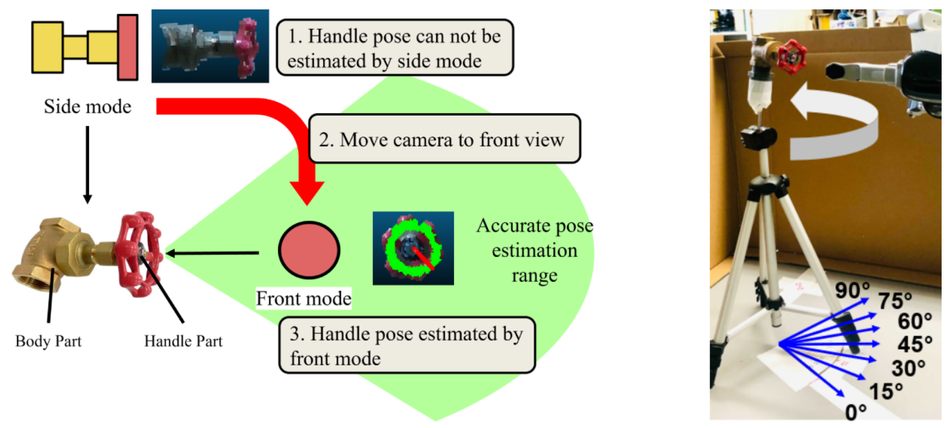

Figure 1:Left: Schematic diagram, Right : Experiment setup

System Architecture

- For object recognition, the scene is regionally segmented using a color-based region segmentation to obtain the PC of the valve [1]. The segmented PC matching is performed based on the Color Signature Of Histograms (CSHOT) feature descriptor.
- To determine the valve mode (front or side) the singular value decomposition is used. The front mode pose estimation is obtained from the RANdom SAmple Consensus (RANSAC) algorithm. If it is a side mode, the body part and pose are obtained using the position and pose relationship with the handle part using the pose integration Hough-Voting method.
- According to the mode, the robot arm performs inverse kinematics to perform either rotational motion of the valve or an arm movement.

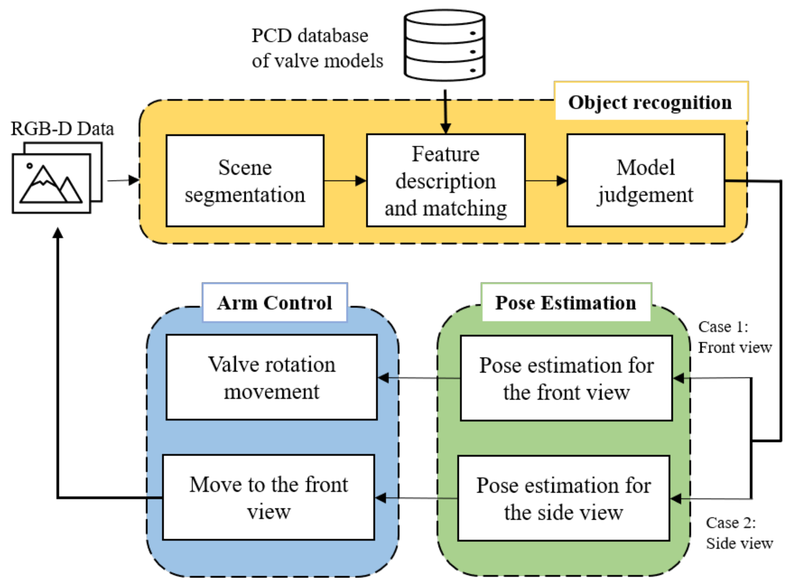

Figure 2:Proposed Object Recognition and Pose Estimation and Arm control model

From the circle from the globe valve normal n, center c, and radius r are computed using three randomly sampled points p. Using the RANSAC algorithm, the cost function (f_h) can be obtained as follows,

$$f_h = (n^T(p-c))^2 + \left(\sqrt{\|p-c\|^2 - (n^T(p-c))^2)} - r\right)^2. \quad (1)$$

Results

Figure 3 shows the results of each step in system architecture.

- Figure 3.(a) Color-based segmentation results in the valve area.
- Figure 3.(b) Feature description using CSHOT.
- Figure 3.(c) Model judgment for front and side modes. Using $\sigma 1$ and $\sigma 2$, the long and short diameters of the valve can be calculated.
- Figure 3.(d) Pose estimation of the front mode and side mode. Figure 3.(d).[1] from RANSAC. Figure 3.(d).[2] From the Hough-Voting method.

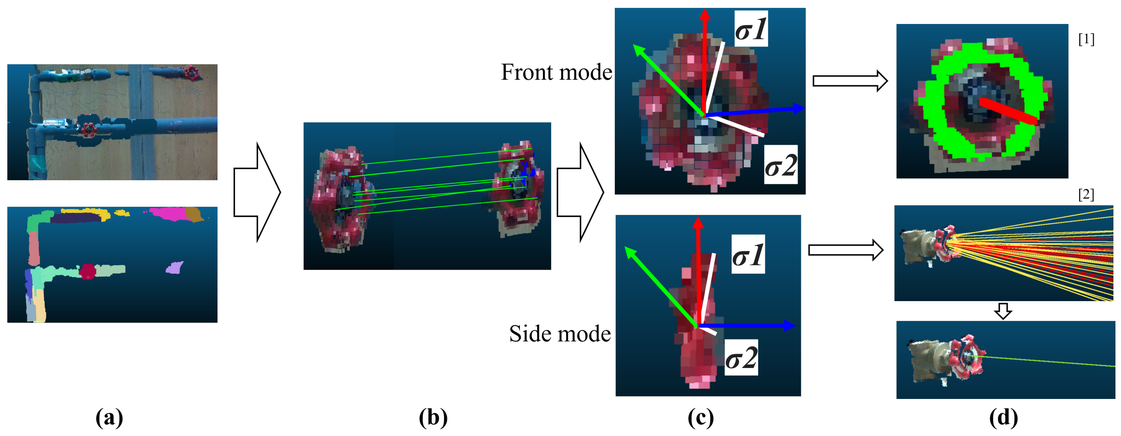

Figure 3:Steps of recognizing and pose estimation of the valve

From figure 4, it is understandable that in the range of 0° to 30°, the recognition accuracy of the valve handle part is high. From figure 5, it is clear that the body part recognizing accuracy is increasing when the angle is increasing as well.

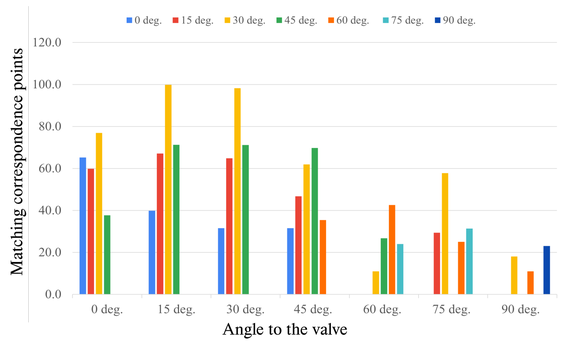

Figure 4:Average model matching correspondence points of the handle part

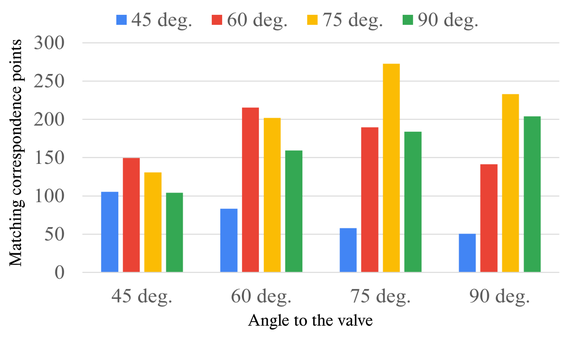

Figure 5:Average model matching correspondence points of the body part

Future Works

- Environment-independent recognition: The color processing of scene segmentation, and feature descriptors may fail due to the lighting level of the environment and observation distance.
- Reduction of execution time: High processing time is required due to the use of many models for matching in the proposed system.
- Generalizing pose estimation: The system is currently working only for the globe valve.

Reference

[1] J. Lim, I. Lee, I. Shim, H. Jung, H. M. Joe, H. Bae, O. Sim, J. Oh, T. Jung, S. Shin, et al. Robot system of drc-hubo+ and control strategy of team kaist in darpa robotics challenge finals. *Journal of Field Robotics*, 34(4):802–829, 2017.

[2] C.-Y. Tsai and S.-H. Tsai. Simultaneous 3d object recognition and pose estimation based on rgb-d images. *IEEE Access*, 6:28859–28869, 2018.