AUTOMATION OF SCREWING NUT BY USING ROBOT ARM WITH HAND-EYE VISION BASED ON DEEP NEURAL NETWORK

5119E023-5 Meng Tianyi Prof. Tetsuya Ogata

Abstract: Recently, collaborative robots are increasingly used on industrial application, especially repeated works with high precision requests and abundantly contact, which is too challenging for human to approach without any directly contacted inner sensorium. However, the conventional mechanical and programming methods present an insufficiency on generalization capability on comparatively complicated motion tasks. In the previous studies, some of researchers performed their works by utilizing Deep Reinforcement Learning[1][2] and demonstrated several enhanced performance on applications, nevertheless, which are too time consuming to establish new system. Therefore, we deployed a deep learning method instead of RL, to propose an ideal balance between system establishment and performance. Moreover, we deployed a hand-eye vision which can observe the contact directly which observes better than [2][3]. Our study is aimed at automation on the task of screwing the nuts on the bolts, which not only achieve a high success rate of approaching procedure, but also reacting an appropriate behavior to various situations of task status. Finally, the result reveals that the model had reached relatively high success rate, but showed an obvious bias of the model.

Keywords: CAE-LSTM, Robotic motion capture, Hand-eye vision

1. INTRODUCTION

Recently, collaborative robots are increasingly used on industrial application, especially repeated works with high precision requests and abundantly contact, which is too challenging for human to approach without any directly contacted inner sensorium.

In actual industrial production applications, most of the robots approached from a conventional mechanical or programming methods. However, traditional solutions are limited to precision and need to improve the performance. It is also greatly time-consuming to adjust parameters while it be deployed into a new productive environment. Therefore, we focused our sights on deep neural networks(DNN), which approach targets nonlinearly, solve more problems that traditional machine learning method .

The purpose of our study is to make a progress on the performance of certain task(screw nuts on the bolts) by adding a hand-eye vision.

To achieve this purpose, we approach our goals follows:

1. Estimate that how much the hand-eye vision can contribute to the robot to attain a better performance in the motion capture task(screw nuts on the bolt). We decide to use YOLOv3[4] in this module to test if it can successfully detect tiny items like bolts.

2. Design an end effector with hand-eye vision which can observe contacts directly better than conventional vision[3]. We decided to add a camera inside the nut which can directly observe contacting movements that outer camera cannot capture.

3. Collected data of approaching phase by collecting reversed data while direct teaching, which collect data of the procedure of the end effector getting away from such bolt to improve the quality of dataset.

4. Integrated a system to approach the screw on task. We use DL instead of RL which is too time-consuming[1][2]. In this study, we would train the end effector to navigate to the bolt automatically by using CAE-LSTM[5]. Moreover, since the studies related to post-pegging task procedure are insufficient, we decide to improve the integrity of the system, integrate a module to send appropriate command while the end effector has collided with anything.

2. METHODOLOGIES AND SYSTEM COMPOSITION

The flow chart of the system of autonomous navigation and screwing on nuts is indicated in Figure 1.

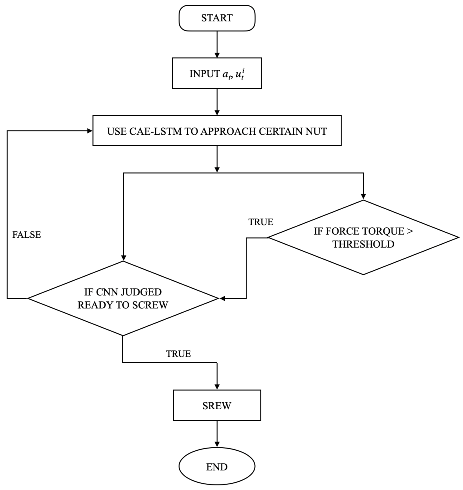

Figure 1. The flow chart of the system.

In the autonomous approaching phase, we decided to train sequence data which contains joint angles and image features to make the universal robot to learn proper behaviors to approach certain bolt. The end effector is designed to be able to pick up the nut, measure force torque and visualize bolt with a n inner camera inside(Figure 2).

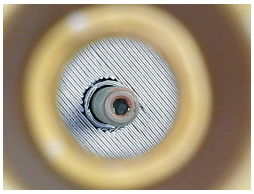

Figure 2. The view of the camera inside the catcher

The architecture of system of CAE-LSTM part is illustrated in Figure 3.

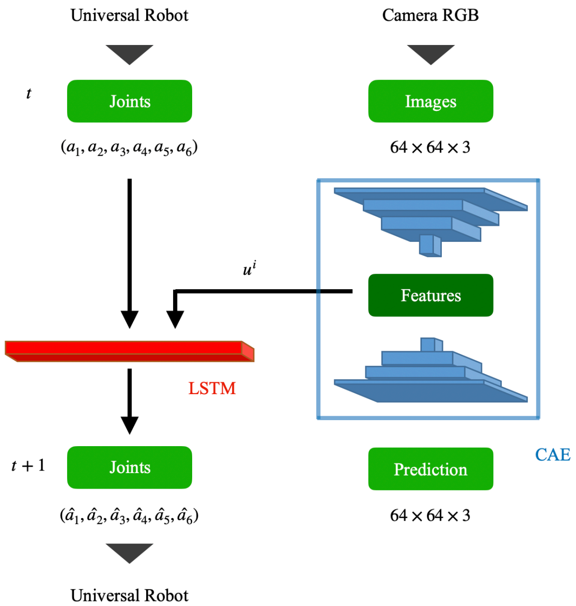

Figure 3 The architecture of CAE-LSTM.

And since the final purpose of our experiment is to screw on the nut to certain bolt, we need to keep tracking if the end effector is on a position that feasible to screw on the nut. Thus, we keep predicting realtime captures by using a CNN model.

3. EXPERIMENT

We decided to train the model by direct teaching. The data collection, we decided to collect data of the procedure of the end effector getting away from such bolt in order to gain a smooth path.

We collected 600 sets of data chose 27 sets of data in (9×3) directions and 5 set of data in (5×1)directions.

And for the evaluation, we estimate the model in offline and online prediction. In offline prediction, we estimate the model by calculating loss of joint angles and compared predicted images with ground truth images. In online prediction, we calculate success rate statistics in 150 frames which covers 5 directions of test data.

4. RESULTS

In the offline prediction, the joint angle loss approximately converge to 0(Figure 4).

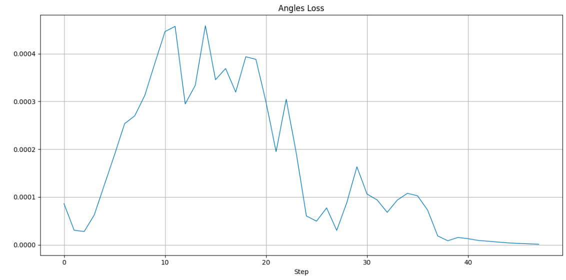

Figure 4. The angle loss of offline prediction.

And in the online prediction, the average success rate is 0.92(Figure 5), which addressed that the model can complete the controlling task. Moreover, right and below presented a lower capability of prediction. We assumed that it is because of the difficulties of pulling the robot while data collection, which deteriorated the data from those two directions.

LEFT-TOP	TOP 1.0	RIGHT-TOP
LEFT 1.0	CENTER 1.0	RIGHT 0.9
LEFT-BELOW	BELOW 0.7	RIGHT-BELOW

Figure 5. The success rate of online prediction.

5. CONCLUSION

We designed an end effector with camera inside the nut, proposed an efficient method of data collection of direct teaching, and trained a model with high performance of the certain approaching and screwing task. It presented a low loss in offline prediction and a high success rate in online prediction. The results of online prediction showed a relationship between data's quality and success rate in different directions.

REFERENCE

[1] T. Inoue, et al. Deep Reinforcement Learning for High Precision Assembly Tasks. arXiv:1708.04033v2, 2017.

[2] D. Wang, et, al. Policy learning in SE(3) action spaces. arXiv:2010.02798v2, 2020.

[3] A. Ishak, et, al. Eye in hand robot arm based automated object grasping system. Periodicals of Engineering and Natural Sciences, Vol. 7, No. 2, pp.555-566, 2019.

[4] J. Redmon, et al. YOLOv3: An Incremental Improvement. arXiv:1804.02767, 2018.

[5] S. Hochreiter, et al. Long Short-Term Memory. Neutral Computation 9(8): 1735-1780, 1997.