Compliance and Learning for Robust In-Hand Manipulation of Various Sized and Shaped Objects

Takashi SATO, Alexander SCHMITZ, Satoshi FUNABASHI, Sophon SOMLOR and Shigeki SUGANO, Fellow, IEEE

Abstract—Moving objects within the hand is challenging, especially if the objects are of various shape and size. In this paper we use machine learning to learn in-hand manipulation of such various sized and shaped objects. The TWENDY-ONE hand is used, which has various properties that make it well suited for in-hand manipulation: a high number of actuated joints, passive degrees of freedom and soft skin, six-axis force/torque (F/T) sensors in each fingertip, and distributed tactile sensors in the skin. A dataglove is used to gather training samples for teaching the required behavior. The object size information is extracted from the initial grasping posture. After training a neural network, the robot is able to manipulate objects of untrained sizes and shape. Compared to interpolation control, the adaptability for the objects of untrained sizes and shape is greatly extended. In particular, analytical solutions are not exactly known beforehand, and it would be preferential if the desired behavior could be trained instead of programmed.

I. INTRODUCTION

After grasping an object, often in-hand manipulation is necessary before the actual task can be performed. For example, after picking up a pen it is necessary to set the pen in the appropriate position before beginning to write. The necessary change of grasping posture is challenging without an additional support. In particular, analytical solutions are difficult to achieve as the shape and size of the object is usually not exactly known beforehand, and the hand and the object might include flexible materials, which are hard to model. Moreover, the current touch state has to be taken into account, and it would be preferential if the desired behavior could be trained instead of programmed.

In previous research the intrinsic compliance (due to springs and soft skin) of the TWENDY-ONE hand was exploited to enable in-hand manipulation by simple interpolation control (each joint angle is changed simultaneously from the initial grasp posture to the final grasp posture), but because no tactile sensor feedback was used, the in-hand manipulation was unstable and the final grasp depended on the initial grasp. With machine learning the in-hand object manipulation possibilities could be extended, but the learning was specific to an object of a certain size and shape. In general, even though a growing number of research on in-hand manipulation exists, stable in-hand manipulation for various sized and shaped objects remains an open research problem. In the current paper we will address this deficiency and enable the robot to move objects of various sizes and shapes within its hand. In order to achieve this, the object size is extracted from the initial grasping posture, and the desired motion is learned from examples provided with teleoperation via a dataglove.

II. ROBOTIC SYSTEM

The hand of the human symbiotic robot TWENDY-ONE, has 16 DOF, as depicted in Fig. 1. The DIP and PIP joints of the index, middle and little finger are coupled, and the hand is actuated by 13 small electric motors integrated in the joints. The DIP and MP1 joints also include springs; there are no springs for the thumb. For the joints with springs, the actual joint angles can be calculated as the motor angles minus the spring displacements. The hand is also covered with a soft skin with 241 distributed tactile skin sensors for the whole hand. In addition, 6-axis F/T sensors are included in each fingertip. The hand is about 20 cm long and the palm is 10 cm wide. For the current paper, only the thumb and the index finger were used, and from the distributed tactile sensors only the ones in the fingertips. Consequently, the 3 motors of the index finger and the 4 motors of the thumb were controlled.

Figure 1. The hand of the human symbiotic robot TWENDY-ONE.

III. METHOD AND RESULTS

The goal task is to learn to move objects of various sizes and shapes with the thumb and index finger from the bottom of the index finger to its side. This task was chosen due to its high difficulty. A feedforward neural network was used to learn the next motor angles depending on the current grasping state as well as the object size and shape. To implement the network, Theano was used, which was later used also for deep learning. The input for the neural network is as follows: 7 motor angles, 2 spring displacements, 2*6 measurements from

Satoshi Funabashi, Alexander Schmitz, Takashi Sato and Sophon Somlor are with the Sugano Lab, School of Creative Science and Engineering, Waseda University, Okubo 2-4-12, Shinjuku, Tokyo, 169-0072, Japan (e-mail: schmitz@aoni.waseda.jp).

Shigeki Sugano is with the Department of Modern Mechanical Engineering, School of Creative Science and Engineering, Waseda University, Okubo 3-4-1, Shinjuku, Tokyo, 169-8555, Japan, (phone: 81-3-5286-3264, Fax: 81-3-5272-0948, e-mail: sugano@waseda.jp).
the 6-axis F/T sensors in the fingertips, 2*36 sensor measurements of the distributed sensors in the fingertip skin, 1 object size, and 1 object shape. In total the network input has a dimension of 95. In order to gather training data, the desired behavior was performed through teleoperation with spheres and cylinders of diameter 20, 40, and 60mm. In this way, 300 trials (2 shapes * 3 diameters * 50 trials per object) of successful in-hand manipulation were recorded. Each of the 300 trials was divided into 50 time steps, resulting in 15000 datasets. The values of all sensor measurements as well as the object parameters (size and shape) were normalized to values between −1 and 1. The object size was not provided, but instead was calculated using a kinematic model of the hand and the sensor measurements (motor angles and spring displacements) of the initial grasp. It was later discovered that it is not necessary to provide the object shape to the network, and therefore it was obmitted. On the other hand, experiments showed the importance of size and tactile information.

The method enabled the robot to generate very robust in-hand manipulation for objects of sizes and shapes that it had not been trained with, see Fig. 3 and 4. Compared to simple interpolation control, the adaptability for the initial grasping posture could be largely extended, see Fig. 5. When using a deep neural network, the number of required training sets could be decreased: only 900 datasets were required for the supervised training. In summary, compared to prior results, we have extended the in-hand manipulation capabilities of the robot.

ACKNOWLEDGMENT

Part of this research was supported by the JSPS Grant-in-Aid for Scientific Research (S) No. 25220005. Part of this research was supported by Research Institute for Science and Engineering, Waseda University.

Figure 2. The neural network used for learning the behaviour.

Figure 3. An untrained, egg-like shaped object was also used to evaluate the robustness of the learned in-hand manipulation skill. This figure shows the different sizes and poses that were used for evaluation.

Figure 4. In-hand manipulation of untrained objects. Top: egg, diameter 40, pose B. Middle: egg, diameter 40, pose C. Bottom: egg, diameter 50, pose A.

Figure 5. The maximum initial displacements of the spheres of different diameter for which successful in-hand manipulation could be achieved is shown.

Figure 6. Example of an in-hand manipulation for a sphere of diameter 30mm. Even though the initial grasping posture is out of the center of the fingertips (x = −10mm, y = −10mm) the handling was successful.