Towards Complex and Continuous Manipulation: A Gesture Based Anthropomorphic Robotic Hand Design

Li Tian, Hanhui Li, Qifa Wang, Xuezeng Du, Jialin Tao, Jordan Sia Chong, Nadia Magnenat Thalmann, and Jianmin Zheng

Abstract—Most current anthropomorphic robotic hands can realize part of the human hand functions, particularly for object grasping. However, due to the complexity of the human hand, few current designs target at daily object manipulations, even for simple actions like rotating a pen. To tackle this problem, we introduce a gesture based framework, which adopts the widely-used 33 grasping gestures of Feix as the bases for hand design and implementation of manipulation. In the proposed framework, we first measure the motion ranges of human fingers for each gesture, and based on the results, we propose a simple yet dexterous robotic hand design with 13 degrees of freedom. Furthermore, we adopt a frame interpolation based method, in which we consider the base gestures as the key frames to represent a manipulation task, and use the simple linear interpolation strategy to accomplish the manipulation. To demonstrate the effectiveness of our framework, we define a three-level benchmark, which includes not only 62 test gestures from previous research, but also multiple complex and continuous actions. Experimental results on this benchmark validate the performance of current methods is promising. Nevertheless, few current methods can tackle the problem of daily object manipulations, even for a simple action like spinning a pen. This is because they are highly specialized for their particular purposes and lose the adaptability. For example, a soft robotic hand targeting at object grasping might not realize the functions of abduction and adduction for all fingers. This problem should not be ignored, as manipulating common objects in a human-like way, is one of the basic requirements for social robots and humanoids.

I. INTRODUCTION

Artificial hands remain one of the hardest problems in robotics [1], due to the lack of comprehensive understanding on the actuation and sensory systems of the human hand. Earlier studies [2], [3], [4] try to fully replicate functions of the human hand via complicated mechanical structures and actuation systems. Although these robotic hands have ranges of motion (ROM) and degrees of freedom (DOF) similar to those of human hands, or even have the ability to complete astounding manipulation tasks like solving a Rubik’s cube [4], they are costly to fabricate and their dexterity can still be improved.

Therefore, recent robotic hand designs focus more on simplifying mechanical structures and realizing partial functions of human hands. To the best of our knowledge, there are two notable development directions for such a purpose. The first direction is based on the concept of anatomically correct robotic hand [5], [6], [7], [8], which imitates the critical biomechanical structures of the human hand, such as bones, joints and tendons. This strategy does lower the complexity of the designing process, yet maintaining the balance between dexterity and complexity remains a challenge. For instance, in one of the latest anatomically correct robotic hand [8], 30 servo motors are used to realize 33 grasping gestures of Feix [9] and the Kapandji test [10].

The other trend includes the soft robotic hands that utilize soft materials and actuators, such as gas/liquid pressure [12], [13], [14] and artificial muscles [15], [16], [17]. Robotic hands of this type have remarkable deformability, and hence they can accomplish the tasks of object grasping safely [18], [19], [20], [21]. However, their mechanical structures and kinematic models are different from that of the human hand significantly. Hence it is difficult for them to realize natural and human-like manipulations.

In our opinion, for the problem of object grasping, the performance of current methods is promising. Nevertheless, few current methods can tackle the problem of daily object manipulations, even for a simple action like spinning a pen. This is because they are highly specialized for their particular purposes and lose the adaptability. For example, a soft robotic hand targeting at object grasping might not realize the functions of abduction and adduction for all fingers. This problem should not be ignored, as manipulating common objects in a human-like way, is one of the basic requirements for social robots and humanoids.

To address the above issues, we propose a novel framework, in which the design of robotic hands and object manipulations depend on a set of grasping gestures. The intuition behind our framework is that, most daily object manipulation tasks can be represented by the set of grasping gestures. Such a strategy is possible, as the pioneer research by Bullock et al. [22] has tried to classify common manipulation tasks at a coarse level. Meanwhile, we can measure the ROM of each base gesture, so that our robotic hand design can be simplified to provide the only necessary ROMs. In this way, the proposed design can achieve the excellent balance between simplicity and dexterity. Besides, since there are extensive studies on achieving various grasping gestures, we can avoid building our solution from scratch, while current methods can be augmented with our framework to tackle the object manipulation problem.

Consequently, based on the proposed framework, we design a 13-DOF anthropomorphic hand, which is much simpler yet more versatile compared with previous methods [8], [23]. Besides, we propose a key-frame interpolation method, in which we select multiple base gestures as the

All the authors are with Nanyang Technological University, Singapore. This research is supported by the National Research Foundation, Singapore under its International Research Centres in Singapore Funding Initiative. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of National Research Foundation, Singapore. Hanhui Li is the corresponding author (hanhui.li@ntu.edu.sg).
Gesture based Design

33 Feix Grasping Gestures

Simplified Design

measure ROM of 33 gestures

key frames, and then adopt the simple linear interpolation to control the process of manipulating an object. To fully evaluate the dexterity of our design, we further introduce the complex and continuous manipulation benchmark, which has tasks of three difficulty levels.

In summary, this paper makes noteworthy contributions to anthropomorphic robotic hand design as follows:

• We introduce a novel framework to provide previous grasping-oriented methods with the adaptability to object manipulations, which also sheds light on designing highly dexterous robotic hands.

• We collect ROMs of a set of base gestures, which are valuable data for robotic hand design and algorithm development for object manipulation.

• We propose a simplified dexterous hand model of 13 DOFs and a key-frame interpolation method, which are validated on our three-level benchmark for object grasping and manipulation.

II. THE GESTURE BASED FRAMEWORK

In this section, we introduce the proposed gesture based framework in detail. As demonstrated in Fig. 1, we first select a set of grasping gestures as the bases. We then measure the ROM of the human hand for each gesture, so that we can design a robotic hand with the necessary dexterity to accomplish the base gestures. Finally, given a manipulation task, we choose multiple gestures as the key frames, and adopt the interpolation method to control our robotic hand.

Due to the complexity of the human hand anatomy, we summarize the terms and abbreviations related to the human hand structures in Table I for easy understanding and reading.

TABLE I

| Abbreviation | Definition |
|--------------|------------|
| Abd/Add      | abduction/adduction |
| DIP          | Distal interphalangeal |
| PIP          | Proximal interphalangeal |
| MCP          | Metacarpal |
| CMC          | Carpometacarpal |
| LUM          | Lumbrical muscles |
| DOF          | Degree of freedom |
| ROM          | Range of motion |
| CCM          | Complex and continuous manipulation |

A. The Base Gestures

Selecting the base gestures allows us to understand the essential functions that need to be implemented. In this paper, we choose the grasp taxonomy defined by Feix et al. [9], which has organized the common human hand configurations into 33 gestures. Any other gesture can be included but we find that these 33 gestures are adequate for various in-hand manipulation tasks.

Given the selected base gestures, we consider the human hand as the template for measuring the ROMs. Previous studies [24], [14] suggest that each finger can be characterized with a kinematic model of 3 DOFs (3 joints with flexion and extension only). Hence, for each gesture, we annotate 15 landmarks manually from two views, as demonstrated in Fig. 2. We use the angles of joint rotation in a fixed \( xyz \) coordinate system as the metric of ROM. In order to obtain a precise template, besides flexion and extension, we also take abduction and adduction into account, and hence 4 joint
rotation angles will be recorded for each finger. Each gesture is measured 3 times, and we use the average rotation angles as the final representation of the gesture.

B. Design of The Proposed Hand

The base gestures can support us to simplify the robotic hand design significantly. For example, we notice that in the abduction and adduction motions of the ring finger, the joint angle is in $[-18^\circ, 6^\circ]$ for all grasping gestures, which is ignorable compared with that of other fingers. Therefore, we can safely remove the tendons for controlling the abduction and adduction of the ring finger.

Specifically, our robotic hand design is a simplified system with 13 DOFs, as demonstrated in Fig. 3. Following the intra-finger constraint proposed in [25] that the distal interphalangeal (DIP) and proximal interphalangeal (PIP) joints are always bent together, we use 1 Bowden cable to control these 2 joints simultaneously. Besides, since we do not need to implement the abduction and adduction motion of the ring finger, we further simplify the design by using 1 Bowden cable to control the DIP, PIP and metacarpal (MCP) joints of the ring finger together. For the rest 3 MCP joints, 2 Bowden cables are attached to each of them to realize the motions of flexion/extension and adduction/abduction. The same mechanism is applied to the carpometacarpal (CMC) joint of the thumb.

Our design has preserved the advantages of anatomically correct robotic hands in resembling the biomechanical structures of the human hand. Moreover, As it needs to meet the minimum ROM requirements of the base grasping gestures, the dexterity of our fabricated hand is guaranteed.

C. Materials and Fabrication

Our design can be fabricated efficiently by leveraging the advantages of 3D printing in fast manufacturing. Based on our previous studies on 3D modeling of hands [26], [27], we construct a printable 3D model consisting of three layers, i.e., the skin, tissues, and bones. Note that the tissue layer is a unique soft deformable structure, so that the 3D model can complete various grasping gestures successfully.

The mechanical structures of the 3D model is shown in Fig. 4. As mentioned above, Bowden cables are used as tendons and muscles to drive the phalanges: For the thumb, the flexor pollicis longus is used to control the flexion and extension of the DIP joint and the PIP joint, while the flexor pollicis brevis for that of the MCP joint. And the adductor pollicis is used for the abduction and adduction of the thumb. For the index finger, the middle finger and the little finger, the flexor digitorum tendons are used for the flexion and extension of the DIP joints and the PIP joints, while the lumbrical muscles (LUM) for that of the MCP joints. The abduction and adduction motions of these finger are controled by the interossei muscles. The structure of the ring finger is omitted, as all its three joints are controled by the same Bowden cable. All the Bowden cables we used in this paper are of 7 strands with a diameter of 0.75 mm, except for those mimicking the LUM, which are of 19 strands with a diameter of 1.25 mm and have a higher modulus of rigidity.

Besides, the flexor pulleys are printed together with the bones to maintain the apposition of tendons and bones. Two
types of plastic tendon sheaths are inserted into the bones to aid the routing of tendons. To connect the bones, we use fishing wires as the collateral ligaments, which has a diameter of 0.34 mm, straight tension of 17 lb and knot tension of 15.5 lb. Articular cartilages of sphere shape are used to reduce the friction between bones.

As to the 3D printing materials, we use the same materials as in [27]. Our actuation system is composed of 13 MG996R servo motors (torque: 11 kg/cm). In Table III, we summarize the configurations of our robotic hand, and compared it with two state-of-the-art methods. From this table, we can see that the complexity of our robotic hand is moderate, and as we will demonstrate in the experimental section, the proposed design is robust enough to handle different grasping and manipulation tasks.

D. Key-frame Interpolation for Object Manipulation

Last but not least, we introduce a simple key-frame interpolation method to demonstrate the potential of the gesture based framework in tackling daily manipulation tasks. More advanced techniques, such as deep reinforcement learning [4], can be applied within the proposed framework, but this is beyond the scope of this paper.

Given a target manipulation, e.g., rotating a pen 180 degrees, we can first represent it by multiple base gestures, which we refer as the key frames. These key frames can be selected manually, or via heuristic strategies, e.g., we can consider gestures as states in a Markov chain and learn to sample optimal states [28]. Here we assume both the initial gesture and the end gesture are palmar pinch (with the index finger and the thumb), and the 4 selected key frames are abduction grip, tripod, palmar pinch (with the middle finger and the thumb), and prismatic 2 finger, as demonstrated in Fig. 1. Without loss of generality, let \( \theta_i \) denote the 20 ROM values of one of the key frames, and \( \theta_{i+1} \) of the next key frame. Then, given an interval of \( T \) frames, we consider the simple linear interpolation method to obtain the ROM of the intermediate gesture at the \( t \)-th frame as follows:

\[
\theta_{i,t} = \theta_i + \frac{t(\theta_{i+1} - \theta_i)}{T},
\]

where \( 1 \leq t \leq T \). Such a method is practical due to the following two reasons: First, the 33 base gestures of Feix are static and stable per se, which indicates that if the interval or \( \theta_{i+1} - \theta_i \) is small enough, the object can be considered to be relatively stable to the robotic hand. Second, within the selection process of the key frames, we can impose heuristic constrains, e.g., the motion of the robotic hand should not be obstructed by the object. In this way, we can decompose various daily manipulation tasks into gestures and perform them with our robotic hand.

III. EXPERIMENT

To validate the dexterity and quality of our gesture based robotic hand, we conduct extensive experiments in this paper. We first measuring multiple kinematic quantities of our robotic hand, and then perform qualitative analysis on a three-level benchmark for object manipulation.

A. Kinematics Analysis

To begin with, we measure the relationship between tendon excursion and joint rotation angle of our robotic hand. As suggested in [29], the tendon excursion \( E \) and the joint angle \( \phi \) of the human finger can be described as the following linear model:

\[
E = r \phi,
\]

where \( r \) is the instantaneous moment arm. This indicates that if the relationship between \( E \) and \( \phi \) of our robotic hand also follows a similar linear function, then our design can replicate the movements of tendons precisely. Hence we record these data of our fabricated index finger for evaluation. The lengths of the distal, middle and proximal phalanges of the index finger are 22.2 mm, 24.2 mm and 41.9 mm, respectively. The results are reported in Fig. 5 and they demonstrate the obvious liner pattern. We further fit these data with the linear regression model (i.e., \( \phi = aE + b \)), and obtain the coefficient of determination \( R^2 \) for all three joints. All \( R^2 \) values are larger than 0.96, which suggest that the linear model fits these data well and hence validate our design.

Following [6], we also adopt the fingertip trajectory as a measure of motions. As presented in Fig. 6, the fingertip trajectories of our robotic hand show that it can reach a large area. These results suggest that our design has the notable flexibility and the potential for completing various gestures.

At last, we summarize the ROMs of the human hand [7] and our robotic hand, as well as our measured ROMs of completing 33 grasping gestures in Table III. These results show that it is not necessary to design a robotic hand

| Evaluation Method | Biomimetic Hand [6] | HR-hand [7] | Ours |
|-------------------|---------------------|-------------|------|
| **Actuator**      | 3D printed, ABS     | Molded, resin | 3D printed, resin |
| **Tendon and Muscle** | 62 gestures and CCM | 15 gestures | 62 gestures and CCM |
| **Driven Cable**  | Soft cable          | Soft cable  | Bowden cable |
| **Ligament**      | Silicone ligament   | Fishing wire| Fishing wire |
| **Tissue and Skin** | Fingertip cap only  | N.A.        | Elastic tissue, silicone skin |
| **Bone**          | 3D printed with bones | N.A.        | N.A. |

2 data from [https://www.verywellhealth.com](https://www.verywellhealth.com) and [https://www.orthopaedicsone.com](https://www.orthopaedicsone.com)
which can fully reach the maximum ROMs of the human hand, since the ROMs of the grasping gestures are actually smaller. Besides, our design can duplicate about 80% of the human hand ROMs and 90% of the grasping ROMs, which guarantees our probability of tackling manipulation tasks.

### Table III

**Comparison of the ROM of human hand and our robotic hand, and that for completing 33 grasping gestures.**

| Finger | Joint | Human | Grasping ROM | Ours |
|--------|-------|-------|--------------|------|
| Thumb  | DIP   | [0, 90] | [0, 70] | [0, 70] |
|        | MCP   | [0, 70] | [0, 70] | [0, 75] |
|        | CMC   | [0, 53] | [0, 48] | [0, 61] |
|        | Abd/Add | [-40, 50] | [0, 50] | [0, 50] |
| Index  | DIP   | [0, 80] | [0, 70] | [0, 70] |
|        | MCP   | [0, 90] | [0, 90] | [-15, 82] |
|        | Abd/Add | [-20, 25] | [-20, 22] | [-8, 26] |
| Middle | DIP   | [0, 80] | [0, 80] | [0, 90] |
|        | MCP   | [0, 90] | [0, 90] | [-7, 95] |
|        | Abd/Add | [-20, 25] | [-20, 20] | [-6, 15] |
| Ring   | DIP   | [0, 80] | [0, 80] | [0, 80] |
|        | MCP   | [0, 90] | [0, 90] | [-5, 95] |
|        | Abd/Add | [-20, 25] | [-18, 6] | - |
| Little | DIP   | [0, 80] | [0, 80] | [0, 69] |
|        | MCP   | [0, 90] | [0, 90] | [-4, 90] |
|        | Abd/Add | [-14, 20] | [-9, 28] | - |

B. Complex and Continuous Manipulation

We perform a series of object grasping and manipulation tasks to demonstrate the effectiveness of the proposed framework. Unlike conventional methods that only utilize single benchmark (e.g., Feix’s grasping gestures) for evaluation, here we introduce the idea of complex and continuous manipulation, which organizes daily manipulation tasks into the following three difficult levels:

- **Level 1: Single Grasping/Manipulation.** Tasks at this level can be tackled by a single gesture, such as the grasping gestures of Feix. As far as we know, most current methods are tested at this level.

- **Level 2: Complex Manipulation.** Tasks at this level require more than one gesture. For example, in our mentioned
Fig. 8. CCM completed by the proposed design. (a) Palmar pinching and rotating a pen. (b) Tripod based rotation of a ping-pong ball. (c) Crawling. (d) flicking and spinning a pen. (e) flicking a balloon. (f) climbing up. (g) rotating the top layer and right layer of a Rubik’s cube, as well as overturning it. Best viewed on a high-resolution screen.

At last, we also record the time and gestures for completing each of the 7 tasks, which could serve as the level 3 baseline for future research. As demonstrated in Fig. 8, most of these tasks are completed within a few seconds. It is noted that the task of rotating a pen consumes 23 seconds, as demonstrated in Fig. 2. This is because we use 5 gestures as the key frames. It is expected that with more advanced learning techniques, fewer gestures might be used and the processing time will be further reduced.

IV. CONCLUSION

In this paper, we propose the gesture based framework, of which the core is the set of base gestures and their ROMs. Given the widely-used gestures as the bases, we devise the simplified 13-DOF robotic hand to provide the essential functions for achieving them. We also demonstrate that complex manipulation tasks can be subdivided into multiple base gestures, and we adopt the key-frame interpolation method to complete the tasks. The proposed hand is evaluated on our CCM benchmark and demonstrates its remarkable dexterity.

In all, we hope that the gesture based framework, as well as the three layer CCM benchmark can provide a considerable baseline for future studies.

REFERENCES

[1] C. Piazza, G. Grioli, M. Catalano, and A. Bicchi, “A century of robotic hands,” Annual Review of Control, Robotics, and Autonomous Systems, vol. 2, pp. 1–32, 2019.
[2] C. Lovchik and M. A. Diftler, “The robonaut hand: A dexterous robot hand for space,” in IEEE International Conference on Robotics and Automation, vol. 2, 1999, pp. 907–912.
[3] F. Rohling, R. Haschke, J. J. Steil, and H. Ritter, “Platform portable anthropomorphic grasping with the bielefeld 20-dof shadow and 9-dof tum hand,” in IEEE/RSJ International Conference on Intelligent Robots and Systems, 2007, pp. 2951–2956.
[4] I. Akkaya, M. Andrychowicz, M. Chociej, M. Litwin, B. McGrew, A. Petron, A. Paino, M. Plappert, G. Powell, R. Ribas et al., “Solving rubik’s cube with a robot hand,” arXiv:1910.07113, 2019.
[5] A. D. Deshpande, Z. Xu, M. J. V. Weghe, B. H. Brown, J. Ko, L. Y. Chang, D. D. Wilkinson, S. M. Bidic, and Y. Matsuoka, “Mechanisms of the anatomically correct testbed hand,” IEEE/ASME Transactions on Mechatronics, vol. 18, no. 1, pp. 238–250, 2011.
[6] Z. Xu and E. Todorov, “Design of a highly biomimetic anthropomorphic robotic hand towards artificial limb regeneration,” in IEEE International Conference on Robotics and Automation, 2016, pp. 3485–3492.
[7] A. M. Faudzi, J. Ooga, T. Goto, M. Takeichi, and K. Suzumori, “Index finger of a human-like robotic hand using thin soft muscles,” IEEE Robotics and Automation Letters, vol. 3, no. 1, pp. 92–99, 2017.
Fig. 9. Displacements of joints during rotating a pen. Pose 1-5 are Palmar Pinch, Adduction Grip, Tripod, Palmar Pinch and Prismatic 2 Finger.