Vision-based Belt Manipulation by Humanoid Robot

Yili Qin^{1,2}, Adrien Escande¹, Arnaud Tanguy¹ and Eiichi Yoshida^{1,2}

Abstract—Deformable objects are very common around us in our daily life. Because they have infinitely many degrees of freedom, they present a challenging problem in robotics. Inspired by practical industrial applications, we present in this paper our research on using a humanoid robot to take a long, thin and flexible belt out of a bobbin and pick up the bending part of the belt from the ground. By proposing a novel non-prehensile manipulation strategy "scraping" which utilizes the friction between the gripper and the surface of the belt, efficient manipulation can be achieved. In addition, a 3D shape detection algorithm for deformable objects is used during manipulation process. By integrating the novel "scraping" motion and the shape detection algorithm into our multi-objective QP-based controller, we show experimentally humanoid robots can complete this complex task.

I. INTRODUCTION

Deformable objects, such as cables, cloths, sponges are very common around us in our daily life. Different from rigid objects, deformable objects have configurations of infinite dimension. Their shape can be easily changed by the external forces. In general, their physical properties are difficult to obtain. Predicting their deformation by using traditional modeling-based methods is difficult. For these reasons, many tasks for handling deformable objects in the factory are still done manually by human workers.

Research on humanoid robotics has been going on for several decades [2]. Early research about humanoids mainly focused on the balance and walking. Due to the increasing maturity of technologies, many projects began in recent years have looked at using humanoid robots for industrial production. As part of the Horizon H2020 program in Europe, the COMANOID project [6] aimed at deploying humanoid robots to achieve non-added value tasks in aircraft manufacturing operations. In such cluttered factory environments, humanoid robots can overcome their biped instability and improve their flexibility by adding or reducing contact points with surroundings [14].

In manufacturing, to facilitate their transportation and handling, long deformable materials are coiled round-byround, or layer by layer on bobbins (see Fig. 1 for examples). Once the bobbin is brought at its designated place in the production chain, the next step for workers to handle the material is usually to find their exposed ends, and take it out of bobbins, to then feed the production.

This paper is one of the steps to the bigger goal of having humanoid robots to help human workers to complete the



Fig. 1. Upper left: A long cable is coiled in a wood bobbin; Lower left: In a factory, a long and flat rubber belt is coiled in a steel bobbin; Right: A section of thin and flexible belt is taken out of the bobbin by humanoid robot HRP-2Kai. (For confidentiality reasons, we need to blur the sides of the bobbin we are using in this paper.)

tedious and low-added value tasks. We consider here the case of taking belts out of bobbins. To achieve this goal, we are facing some difficulties:

- It is hard to find the exposed ends of the objects for a robot because there is no special features or markers on the objects. In our case, the manipulated object is a thin and flexible belt, the upper layer belt is laying on the lower layer, this brings great challenges for directly detecting and handling the end of the belt. In some cases, to get the position of the end may be even not feasible for a robot.
- When the robot is handling the belt, knowing the state of the belt is very important and challenging. A wrong estimation in the state of the deformable object can easily cause the operation to fail.
- A humanoid robot is an underactuated system on a floating base. While handling long and flexible objects, controlling balance becomes a very important issue.

In this paper, by utilizing the state-of-the-art technologies and proposing novel methods, we try to resolve these issues. The paper is organized as follows: In section II, we introduce the related works. Section III presents the work about shape detection for the belt. Section IV presents separating toplayer belt by a kind of force-based manipulation, "scraping", which is inspired by human motions. Section V introduces how the robot picks up the bending belt on the ground under the help of visual information. Section VI introduces the experiments and the results.

¹CNRS-AIST JRL (Joint Robotics Laboratory), IRL, Tsukuba, Japan. Corresponding author: Y. Qin s1930205@s.tsukuba.ac.jp

²Department of intelligent and mechanical interaction system, Graduate school of science and technology, University of Tsukuba, Tsukuba, Japan.

II. RELATED WORK

To handle a long and flexible object, a very common strategy is to directly grasp it. Force closure on the contacts improves the controllability and the stability for the objects. Related research has been extensively conducted. Based on grasping, regrasping and locomotion primitive actions, Oin et al. proposed a multi-layer framework for a humanoid robot to install a long flexible cable to two clamps [12]. In many cases, limited by surrounding environments and configurations of a robot, direct grasping is difficult, even not possible. Researches about picking up a thin object can be found in [8] [16]. But this kind of manipulation needs special mechanism design, precisely sensing and handling for the edge of the object. Another idea is to take some preprocessing stages with non-prehensile manipulation, moving the object to a position that is easier to grasp, then perform the grasping. In [4], Hang et al. proposed a planner based on pre-grasp manipulation for picking up a thin object from the table. In some cases like fast manipulation, comparing to prehensile manipulation, non-prehensile manipulation can provide more effective solutions for the robot.

Visual information can significantly improve the effectiveness of robotic manipulation. Since shape of deformable objects can be easily changed during manipulation, being able to know the shape information becomes a very critical issue. Zhu et al. proposed a vision servo approach to deform a cable to desired shapes in 2D plane with a dual-arm robot [17]. Li et al. used "top-sliding" motion to actively deform a flexible printed circuit board and proposed a visionbased control scheme was for the soldering system [7]. By encoding the state of the deformable objects, Hu et al. proposed a deformable object manipulation controller for a robot to bring several kinds of objects to desired shapes [5].

Our contributions are as follows:

- We propose and analyze an efficient force-based manipulation strategy "scraping" for handling a long, thin and flexible belt.
- We resolve a challenging manipulation task by integrating multi-objective Quadratic Programming (QP) based controller [11] [14], deformable object shape detection and "scraping" motion.
- We experimentally verify the approach can be used in the process for belt manipulation by a humanoid robot, showing the approach has the potential to be applied to practical applications.

III. SHAPE DETECTION

The shape of a deformable object in 3D space is a continuous mathematical object that needs to be represented with a finite number of parameters. Some commonly used shape representation approaches are geometric-based method, parametric equation and so on. Considering the data type we get from the robot vision sensor, we choose to use a geometric model composed of discrete points to represent the continuous shape. Thus, a shape detection problem can be simplified and formalized as a point set registration problem. Assume we have two point sets $\mathbf{X} = \{x_1, x_2, ..., x_N\} \in \mathbb{R}^{N \times D}$, $\mathbf{Y} = \{y_1, y_2, ..., y_M\} \in \mathbb{R}^{M \times D}$, N and M are the number of points in each set, D is the dimension of point. Function f is the transformation model. The goal of this problem is to find the optimal transformation f^* , such that point set \mathbf{X} is best aligned to \mathbf{Y} :

$$f^{\star} = \underset{f \in \mathcal{F}}{\operatorname{arg.\,min.}} \quad dist(f(\mathbf{X}), \mathbf{Y}) \tag{1}$$

where \mathcal{F} denotes all possible transformation models for **X** and *dist* denotes the distance function for a pair of points (usually, the Euclidean distance). Figure 2 provides an example for the registration problem.



Fig. 2. Illustration of 3D space point set registration: Point set X (blue) is transformed to align point set Y (red).

Coherent Point Drift (CPD) algorithm proposed by Myronenko et al. [9] is a probability-based point set registration algorithm and is suitable for non-rigid objects. This method takes a Gaussian Mixture Model (GMM) point of view, where each point in \mathbf{X} is regarded as a Gaussian centroid, and the point set \mathbf{Y} is regarded as the samples from the mixture Gaussians. To prevent data from overfitting, coherent point drift regulation is introduced to add constraints between points in \mathbf{X} , and topological structure of the point set can be preserved. This obviously improves the robustness for outliers and occlusions. Tang et al. [13] applied this algorithm for robotic manipulating a rope. Taking their research as a reference, and with only a few minor changes to the initial structure of \mathbf{X} to extend to 2D objects, this algorithm can be used in our research for belt shape detection.

Flow chart for shape detection is shown in Fig. 3. From the RGB-D camera, we get the image data and the point cloud data. Voxel-based filter is then used to reduce the density of the point cloud. A sparse point cloud is thus generated. To separate the belt's point cloud, a color-based segmentation method is used in the image space. Here, to improve the robustness to ambient lighting, we choose HSV space based segmentation algorithm. This gives us a target point cloud of the belt after segmentation. Note that this point cloud would contain noises and part of the point cloud may be lost in some cases. We regard this point cloud as point set **Y**. A manually-defined geometric model with 2D plane structure is then generated. We denote the points in the model as point set **X**. CPD algorithm is implemented to deform the geometric model to best align the belt's point cloud.



Fig. 3. Process of belt shape detection. The color of the points in the geometric model is only for visualization and has no special meaning.

IV. SCRAP AND SEPARATE THE TOP LAYER OF THE BELT BY COMPLIANCE FORCE CONTROL

A. Non-prehensile Manipulation with Friction

As mentioned before, the belt is coiled layer-by-layer in the bobbin. To take the belt out of the bobbin, one strategy is to directly find and grasp the exposed end. Since the belt is thin and flexible, the upper layer is closely lying on the lower layer. This greatly increases the difficulties for the robot to perform without a specially customized gripper (gripper of HRP-2Kai is shown in Fig. 9). Inspired by human manipulating thin objects, we investigate another kind of pregrasp strategy: scraping (Fig. 4). Imitating human behavior, the robot can use its gripper to exert a normal force to the surface of the belt, then keep the normal force and move the gripper along the curve of the bobbin. By utilizing the friction between the gripper and the top layer, the belt will slide with the gripper's movement, and be separated with lower layer. Once the unwound part is long enough, due to the gravity, the end of the belt will be separated from the bobbin, and fall to the ground. Motion scraping is a kind of non-prehensile manipulation. Comparing to prehensile manipulation like grasping, scraping cannot fully control the belt during manipulation, but it has advantages. First, the exact position of the end of the belt can be ignored. Instead, the coiled direction of the belt is enough for robotic manipulation. This will be much easier to specify in a real application scenario. Second, without grasping the belt, meaning that the robot, the belt, the bobbin and the ground do not form a closed-loop structure, motions of the robot become more flexible and safer. Third, without moving to the other side of the bobbin, the robot can unwind the belt, by repeating the scraping motion multiple times. This greatly reduces the time and complexity for practical applications.

To achieve this kind of non-prehensile scraping motion with the robot, friction plays an important role. To simplify the problem, a rough assumption was made for the current research: friction between the gripper and the top layer of the belt, friction between the layers of the belt satisfy the



Fig. 4. Two kinds of belt handling motions demonstrated by human.

following well known static friction formula $F_f = \mu F_N$ The friction force is denoted by F_f . F_N is the normal force exerted on the belt surface. Coefficient of Friction (COF) is denoted by μ .

As the belt is flexible, the normal force F_N on the surface can be transmitted layer-by-layer. We denote the friction force and the COF between the gripper and the belt as F_{f1} and μ_1 , the friction force and COF between the layers of the belt as F_{f2} and μ_2 . To prevent the contact point sliding while scraping, we need to ensure $F_{f1} > F_{f2}$, so as $\mu_1 > \mu_2$. In real situation, the COF between flexible objects is much more complicated and may change according the exerted force, and the sliding between two layers also make the analysis more difficult. But from the simplified analysis above, we got two critical ideas to achieve scraping motion: (i) increase the COF between gripper and belt and (ii) control the exerted normal force on the surface while scraping.

B. Compliance Force Control



Fig. 5. Bobbin-belt model.

Based on the bobbin-belt model in Fig. 5, a circular trajectory along the curve of the bobbin can be computed a priori. While the gripper tracks the trajectory, it also rotates to keep the direction pointing to the center of the bobbin. A compliance force control law is then applied to the gripper. HRP-2Kai is a position-based robot. To directly control the joint torque is not feasible in our case, and it is the same for the exerted force. An F/T sensor is mounted on each wrist for measuring the external forces and torques, so we can implement an indirect control of the exerted force by utilizing the feedback of the sensor. The main idea is to exert the force thanks to the motion of the gripper, then the

measured wrench on the gripper is fed back to the controller to form a close-loop. A classic admittance control law [11] in 3D task space is as follows:

$$\Delta F_{ext} = M\Delta \ddot{x} + D\Delta \dot{x} + K\Delta x \tag{2}$$

$$\Delta F_{ext} = F_{ext}^{mea} + F_{ext}^d \tag{3}$$

$$\Delta T_{ext} = J\Delta\ddot{\phi} + L\Delta\dot{\phi} + C\Delta\phi \tag{4}$$

$$\Delta T_{ext} = T_{ext}^{mea} + T_{ext}^d \tag{5}$$

where translational and rotational displacements are denoted as x and ϕ , measured external force and torque are denoted as F_{mea} and T_{mea} , exerted force and torque are denoted as F_{ext} and T_{ext} , respectively. M, D, K and J, L, C are the gains, while F_{ext}^d and T_{ext}^d are the desired force and torque for this second-order dynamics system.

The stabilizer for humanoid robots proposed by Caron et al. in [1] is used in the whole process of belt manipulation. Implementation of the stabilizer is based on our multiobjective QP-based control framework. A Center of Mass (CoM) task is used by the stabilizer to track the desired CoM position. Two Center of Pressure (CoP) tasks are used for each foot to track the desired wrench. For the torso and the waist, two body orientation tasks are used to regulate their poses. To implement the scraping motion, an admittance task is used to regulate the contact force to the belt, while an endeffector task which controls gripper's position and orientation is used to track the pre-computed circular trajectory. To make these two tasks work together better, different weights are assigned to them while the stiffness and the damping of each task need to be carefully adjusted.

V. PICK UP THE BENDING BELT

With the constraints due to two sides of the bobbin, the piece of belt that has fallen on the ground can naturally appear in two states: flat and bending (Fig. 6). For now, because we empirically observed that the bending state is more likely to occur, we just consider this second situation. Using the visual shape detection, state of the bending part can be easily determined. The height of each part of the bending belt relative to the floor can be determined by simply analyzing the position of all points in the geometric model. Points located at the edge of the geometric model and whose height is greater than the threshold can be used as the grasping points. Position of a pre-grasping point can be determined by adding a certain offset to according grasping point (Fig. 7). Due to the joint limits and limited number of degrees of freedom, picking up an object from the ground is not an easy task for a humanoid robot. To achieve this, the robot need to squat down, lower and shift the CoM to foot corresponding to the gripper used for grasping, and twist its body while grasping the belt (see Fig. 11 (f)-(h)). Such constrained motions can be achieved by whole-body motion control [15], with reduced safety margins for collision avoidance.

VI. EXPERIMENTS AND RESULTS

First, we present the experiments to validate the shape detection algorithm. Then we use the humanoid robot HRP-



Fig. 6. Two possible states of the fallen belt: flat state and bending state.



Fig. 7. Left: Looking at the belt from the view of the robot. White points correspond to the floor, red ones to the belt; Right: Position for grasping. Blue points are the pre-grasping point and grasping point.

2Kai to validate the force-based scraping manipulation strategy and the feasibility of picking up the bending part by visual shape detection.

A. Shape Detection for the Belt

We cut a piece of belt for the experiments. It is 104cm long, 31cm wide and 0.11cm thick, with a scarlet color.

An ASUS Xtion PRO Live camera was used as the vision sensor to monitor the environments. RGB and depth images (640×480) were captured and sent to an Ubuntu 16.04 laptop (Intel i7@2.70GHz + RAM 64GB). Figure 8(a) shows the shape detection results. Just by visual comparison, we can find that the belt's point cloud and the geometric model (which consists of 30 points, 10 rows×3 columns) match well. According [13], the matching error for a rope is less than 2.2*cm* when using a Microsoft Kinect camera. The speed of the algorithm is related to the size of both point sets. In the experiment, after the downsampling process and color segmentation process, the size of the belt's point set is about 400 (which changes with the distance to the camera and the situation of occlusion). Table I lists the execution time when the size of the geometric model changes.

TABLE	Ι	

EXECUTION TIME

Number of points (rows×columns)	Time (ms)
30 (10×3)	170
60 (15×4)	480
120 (20×6)	1800

In practical applications, since the camera is mounted on the head of the robot (see Fig. 9(a)), when the robot handles the belt, part of the belt may be occluded by the body and the bobbin. Considering this, robustness to occlusion is important in choosing a shape detection algorithm. Figure 8(b) shows the results when validating the robustness of the algorithm. While moving the belt, part of the point cloud is lost due to the occlusion of the arm. But the algorithm still performs well.





(b) With occlusion (RGB image and point cloud)

Fig. 8. Shape detection result for the belt.

B. Robotic Manipulation for the Belt

The inside and outside radius of the bobbin are about 36.5cm and 66.5cm, the width between two sides is about 38cm. For the experiments, the right gripper of the robot HRP-2Kai is used to perform the scraping motion. To increase the friction between the gripper and the surface of the belt, several pieces of sponges are attached to the fingers, which is shown in Fig. 9(b). Some assumptions were made before the experiments:

- The bobbin is fixed on the ground and will not move during the experiments.
- The manipulated parts of the belt are inside HRP-2Kai's working space without locomotion.
- The robot is placed at an approximately known position with respect to the bobbin, within a few centimeters and degrees.
- The exposed end of the belt should be fallen on the ground by only one successful scraping manipulation.

The relative positions of the bobbin and the robot were determined experimentally by simulation in Choreonoid [10]. The distance between the bobbin center and the robot is about 95cm. The assumption on the robot position is not limiting: previous works (e.g. [3]) have shown that a robot can walk to a given location using visual servoing. We could use them with a model of the bobbin to reach a correct location.

The target force during the scraping motion is 20N in the direction from the gripper to the center of the bobbin and 0N is the other directions. The target torque is $0N \cdot m$ for each direction of rotation. As we just consider one direction of the force, the admittance gains are $M_z = 2000$, $D_z = 1600$, $K_z = 20$. The forces and torques are shown in Fig. 10.

Because it is difficult to estimate the friction coefficients at play, we determined the target force by trial and error.





(a) Camera on the head

(b) Sponges on the fingers

Fig. 9. The head and the gripper.



Fig. 10. Force and torque of the right gripper during the experiment: Interval I indicates the force when the robot creates a contact with the belt. The controller regulates the contact force to the target value 20N. Interval II indicates the force during scraping motion. Due to the error between the pre-computed circular trajectory and the curve of the bobbin, and the gripper moving quickly along the trajectory, the force varies around the target value. By generating a more accurate trajectory for the gripper, or further adjusting the parameters of the admittance controller, the effect can be improved.

At 10N, the force is too low to slide the top-layer belt and separate it from the lower layer. At 30N, the robot cannot control its balance and falls backward. Better regulation for the CoM of the robot during the manipulation should be able to improve the balance.

The whole experiment takes about 40 seconds. Approaching to the bobbin and regulating the contact force takes about 11 seconds, scraping along the bobbin takes about 7 seconds, detecting the belt takes about 10 seconds (most of the time is spent on manual checking the matching results for safety reasons), picking up the belt takes about 12 seconds. Snapshots of the experiment are shown in Fig. 11.



Fig. 11. Snapshots of the scraping and picking experiment: (a) Initial pose. (b) The robot raises its right gripper and approaches to the belt. (c) Create a contact with the surface of the belt and regulate the force to the target value. (d) "Scraping" motion. (e) Remove the contact and the belt slides to the ground due to the gravity. (f) Detect the shape of the belt and compute the position of grasping point and pre-grasping point. (g) Move the gripper to the position of grasping point and grasp the bending part of the belt. (i) Stand up while picking up the belt. (j) End pose.

VII. CONCLUSIONS AND FUTURE WORK

In this paper, we presented our work on control of a humanoid robot to take a long, thin and flexible belt out of a bobbin. Then grasping for the bending part of the belt on the ground and pick it up was achieved. A novel strategy scraping was proposed for manipulating belt-like objects. Visual detection for the shape of the belt improved the capability for handling flexible objects. The feasibility was experimentally verified. In the future research, we will integrate the shape detection approach into the control loop. As a scraping motion is a kind of non-prehensile manipulation, manipulation for belt-like objects based on real-time visual servoing will be an interesting topic. In addition, we will also generalize the framework to bobbins of different sizes and belts of different materials.

REFERENCES

- S. Caron, A. Kheddar, and O. Tempier. Stair climbing stabilization of the hrp-4 humanoid robot using whole-body admittance control. In 2019 International Conference on Robotics and Automation (ICRA), pages 277–283. IEEE, 2019.
- [2] T. Fukuda, P. Dario, and G.-Z. Yang. Humanoid robotics-history, current state of the art, and challenges, 2017.
- [3] M. Garcia, O. Stasse, J.-B. Hayet, C. Dune, C. Esteves, and J.-P. Laumond. Vision-guided motion primitives for humanoid reactive walking: Decoupled versus coupled approaches. *The international journal of robotics research*, 34(4-5):402–419, 2015.
- [4] K. Hang, A. S. Morgan, and A. M. Dollar. Pre-grasp sliding manipulation of thin objects using soft, compliant, or underactuated hands. *IEEE Robotics and Automation Letters*, 4(2):662–669, 2019.
- [5] Z. Hu, T. Han, P. Sun, J. Pan, and D. Manocha. 3-d deformable object manipulation using deep neural networks. *IEEE Robotics and Automation Letters*, 4(4):4255–4261, 2019.
- [6] A. Kheddar, S. Caron, P. Gergondet, A. Comport, A. Tanguy, C. Ott, B. Henze, G. Mesesan, J. Englsberger, M. A. Roa, P.-B. Wieber, F. Chaumette, F. Spindler, G. Oriolo, L. Lanari, A. Escande, K. Chappellet, F. Kanehiro, and P. Rabaté. Humanoid robots in aircraft

manufacturing. *IEEE Robotics and Automation Magazine*, 26(4), Dec. 2019.

- [7] X. Li, X. Su, and Y.-H. Liu. Vision-based robotic manipulation of flexible pcbs. *IEEE/ASME Transactions on Mechatronics*, 23(6):2739– 2749, 2018.
- [8] T. Matsuno, K. Kanada, F. Arai, H. Matsuura, and T. Fukuda. Strategy of picking up thin plate by robot hand using deformation of soft fingertip. In *Proceedings of the 2005 IEEE International Conference* on Robotics and Automation, pages 2326–2331. IEEE, 2005.
- [9] A. Myronenko and X. Song. Point set registration: Coherent point drift. *IEEE transactions on pattern analysis and machine intelligence*, 32(12):2262–2275, 2010.
- [10] S. Nakaoka. Choreonoid: Extensible virtual robot environment built on an integrated gui framework. In 2012 IEEE/SICE International Symposium on System Integration (SII), pages 79–85. IEEE, 2012.
- [11] K. Pfeiffer, A. Escande, and A. Kheddar. Nut fastening with a humanoid robot. In 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 6142–6148. IEEE, 2017.
- [12] Y. Qin, A. Escande, and E. Yoshida. Cable installation by a humanoid integrating dual-arm manipulation and walking. In 2019 IEEE/SICE International Symposium on System Integration (SII), pages 98–103. IEEE, 2019.
- [13] T. Tang, C. Wang, and M. Tomizuka. A framework for manipulating deformable linear objects by coherent point drift. *IEEE Robotics and Automation Letters*, 3(4):3426–3433, 2018.
- [14] J. Vaillant, A. Kheddar, H. Audren, F. Keith, S. Brossette, A. Escande, K. Bouyarmane, K. Kaneko, M. Morisawa, P. Gergondet, et al. Multicontact vertical ladder climbing with an hrp-2 humanoid. *Autonomous Robots*, 40(3):561–580, 2016.
- [15] E. Yoshida, M. Poirier, J.-P. Laumond, O. Kanoun, F. Lamiraux, R. Alami, and K. Yokoi. Whole-body motion planning for pivoting based manipulation by humanoids. In 2008 IEEE International Conference on Robotics and Automation, pages 3181–3186. IEEE, 2008.
- [16] T. Yoshimi, N. Iwata, M. Mizukawa, and Y. Ando. Picking up operation of thin objects by robot arm with two-fingered parallel soft gripper. In 2012 IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO), pages 7–12. IEEE, 2012.
- [17] J. Zhu, B. Navarro, P. Fraisse, A. Crosnier, and A. Cherubini. Dualarm robotic manipulation of flexible cables. In 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 479–484. IEEE, 2018.