Situation-Specific Grasping of Fabrics

Niklas Fiedler, Jonas Wiese, and Jianwei Zhang

I. INTRODUCTION

Robotic manipulation of fabrics and clothing has garnered significant attention in recent years, reflecting the growing interest in automating tasks that involve deformable materials. Many approaches have been developed to address the complexities associated with robotic manipulation of fabrics. These methods range from using computer vision and machine learning algorithms to improve the understanding of fabric characteristics, to developing specialized grippers. Additionally, the integration of tactile sensors and feedback mechanisms has played a key role in enabling robots to grasp and manipulate fabrics more effectively.

Every manipulation task begins with grasping the item that needs to be manipulated. As this is a very common problem, many existing approaches address the task of grasp point detection [1], [2], [3]. However, these approaches do not consider a situation such as the piece lying flat on a surface without offering a distinct grasp point. This was addressed using a specialized gripper using its mechanical structure to grasp the pieces [4], but this requires specialized hardware which would not be usable for other tasks at the same time.

The specific goals of the manipulation, which may include tasks such as folding, sorting, dressing, or organizing, alongside the context in which the fabric is situated, presents unique challenges. For instance, a fabric lying flat on a table might require a different grasping technique than a piece of clothing that is crumpled within a heap. These variations necessitate a customized approach to reliably and precisely grasp different types of fabrics in diverse scenarios.

Our work addresses these challenges by implementing a pipeline that determines an appropriate grasp required for effective manipulation in a given situation. We demonstrate the approaches in a human-robot interaction scenario, showcasing how the pipeline responds to varying grasping conditions. By enabling robots to adapt their grasping techniques, we aim to offer a generalized interface for grasping fabrics and garments for further manipulation steps and thus help develop stable approaches to larger tasks.

II. FOLD GRASP POINT DETECTION

In most cases, fabrics and garments to be manipulated throw at least small folds. These are a good option for reliable grasps. To grasp garments at their folds, the optimal folds need to be identified first and grasp poses have to be extracted from them. A challenge posed by highly deformable objects such as fabrics is that their shape might drastically

[1]The authors are with group TAMS, Department of Informatics, University of Hamburg, Germany `niklas.fiedler@uni-hamburg.de`

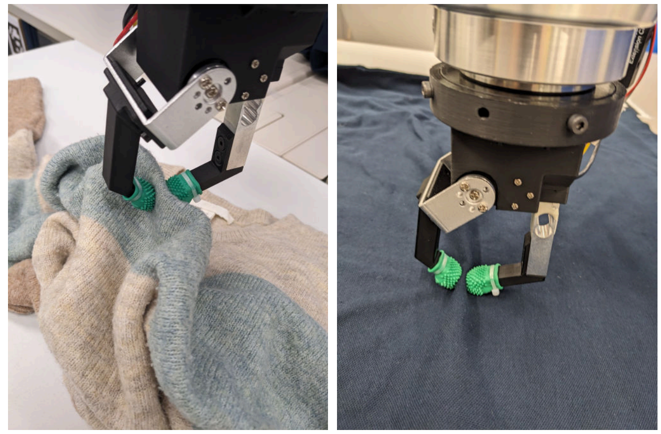

Fig. 1. Images of the two grasp types. **Left**: Grasping an existing fold in the garment. **Right**: Producing a fold to be grasped by pressing down and pinching.

change as soon as the robot starts to manipulate them. In this scenario, that is the case when the robot comes into contact with the piece while approaching a selected grasp pose. Thus, the grasp orientation needs to be adapted to the angle of the fold and care needs to be put into the selection of the height of the grasp point as pushing down on surrounding fabric might alter the fold. We limit the orientation to the Z-axis as all grasps are performed vertically from the top down.

Our fold based grasp point detection uses a point cloud as input. In a preprocessing step the point cloud is transformed into the frame of the table surface and cropped to the area which is to be grasped. We evaluated four approaches of extracting a graspable fold from the point cloud data: (a) Grasping the centroid of the point cloud, (b) grasping the highest point, (c) fitting a cylinder into the point cloud, and grasping the centroid, or (d) the highest point of the points along the surface the cylinder. While we could easily derive the optimal grasping orientation for the cylinders (a finger on each side of the cylinder), a different approach was required for strategy (b). When using the highest point as the grasping point, the surrounding points are analyzed to detect the orientation of the fold by calculating the mean height of a subsection of points on an axis rotated around the grasp point. This process is visualized in Figure 2. When grasping the centroid of the point cloud, a fixed orientation was used. An experiment was conducted by performing 25 grasp attempts with each strategy. The centroid strategy (a) performed worst with 18 successful attempts. Grasping the highest point in the point cloud and grasping the highest point (b) performed best with 25 successful attempts. Both, grasping the centroid (c) and highest point (d) of the cylinder fitted into the point cloud achieved 24 successful trials. How-

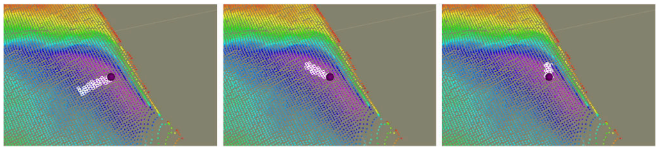

Fig. 2. Visualization of the scanning process used to determine the best grasp orientation. The analyzed point cloud is colored in a gradient along the X-axis. The grasp point is marked in purple and the currently scanned points of the point cloud are white.

ever, it was found that these strategies are highly dependent on good configuration parameters and the process of fitting cylinders into the point cloud causes a sever computational overhead compared to the other approaches. Thus, the highest point of the cropped point cloud in combination with the orientation determined by the scan of the points environment will be used as grasping pose.

III. GRASPING WITHOUT FOLDS

In case no fold is available in the area to be grasped, the robot has to adapt by utilizing a grasping strategy which first creates a fold which is then grasped. This is accomplished by pressing at least two fingers of a gripper down on the fabric while closing it. This pinch motion creates a fold in the material while grasping it as the gripper closes further. As this process relies on the fingertips making stable contact with the material and the material sliding over the table surface, a relatively high friction on the fingertips and low friction on the table surface are required. The downwards pressure of multiple fingertips onto the piece needs to be maintained throughout the grasping procedure to prevent the material from unfolding again. We maintain the downwards pressure by applying a control loop measuring the pressure and dynamically adapting the gripper height above the table. This is especially relevant when using grippers which extend their fingers outwards while closing. Further, the impedance mode of the Agile Robots Diana7 robot arm is utilized to adapt the angle of the gripper to the table surface compensating for unevenness of the table and fabric thickness.

This approach was evaluated in a previous work utilizing a gripper which was not optimized for the task [5]. In the tests performed grasping three different types of fabrics with the same set of parameters, a success rate of 98.7% was reached across 450 grasps. Additionally, 25 grasp attempts were conducted on a flat piece of paper (a new piece was used for each attempt), 19 of which were successful. The main reason for the failed attempts was a lack of friction between the gripper and the piece. This can easily be accounted for by adapting the gripper to the task by applying a rubberized surface to the fingertips. The grasp strategy was also successfully transferred from a semi-parallel gripper to a QB SoftHand to demonstrate its versatility. For this work, the same strategy is applied to another different type of gripper with two fingers, only one of which is actuated as can be seen in Figure 1.

The vertical orientation of the gripper has no significant effect on the grasping result, thus it is fixed. Often, even if there are folds available to grasp, this approach can be used robustly. However, it cannot be applied when multiple pieces lay on top of each other and need to be grasped separately. This scenario occurs for example when sorting a heap of laundry or when picking up garments from a table with a table cloth. In cases like these, the fold based grasp point detection is preferred.

IV. INTERACTIVE GRASP DIRECTION WITH AUTOMATIC SELECTION

To analyze which type of grasp is most appropriate in the given situation, we make use of point cloud measurements of the area to be grasped from. When the fold grasp point detector is able to identify a graspable fold within a designated grasp area, this fold will be grasped using the approach outlined in Section II.

To demonstrate this automatic selection of the appropriate grasp strategy, we set up a demo scenario in which the user points at a piece of fabric indicating the desired grasp point. The scene is captured by a Microsoft Azure Kinect camera. In the RGB image stream of the camera, the Mediapipe hand detector [6] is utilized to detect the tip of the index finger. The point in the image is projected into Cartesian space using the depth measurements taken by the camera. Afterwards, the point is projected down onto the table surface to define the desired grasp point. The fold grasp pose detector is applied onto a small area around the selected point. The grasp pose will be located by selecting the highest point above a preset threshold height above the table. In case it is able to find a graspable fold in the area, it will be grasped as can be seen in Figure 1 **left**. Otherwise, the arm moves to a pose holding the gripper above the garment and applies the fold generating grasping strategy shown in Figure 1 **right**.

V. FUTURE OUTLOOK

With approaches such as diffusion policies, and LLM-based action models processing multimodal input, the capabilities of robots with regard to deformable object manipulation are growing rapidly. As the field continues to evolve, some of these emerging approaches have begun incorporating motion primitives. These motion primitives approach various small actions such as grasping, which can be used to execute more complex tasks. The integration of these motion primitives can be a key element for developing more advanced manipulation strategies.

Invoking a motion primitive with high adaptability and robustness, like our fabric grasping pipeline, can significantly enhance the overall reliability and effectiveness of robotic systems as they tackle larger and more intricate tasks. The adaptability of this grasping pipeline allows it to respond to variations in fabric characteristics and configurations, ensuring that the robot can maintain a stable grip regardless of changes in the environment and across multiple gripper types. This robustness minimizes the likelihood of errors or failures during the manipulation process.

REFERENCES

[1] S. Chitta, E. G. Jones, M. Ciocarlie, and K. Hsiao, "Perception, Planning, and Execution for Mobile Manipulation in Unstructured Environments," *IEEE Robotics & Automation Magazine*, vol. 19, pp. 58–71, Jun. 2012.

[2] A. Ten Pas, M. Gualtieri, K. Saenko, and R. Platt, "Grasp pose detection in point clouds," *The International Journal of Robotics Research*, vol. 36, no. 13-14, pp. 1455–1473, 2017.

[3] H. Liang, X. Ma, S. Li, M. Görner, S. Tang, B. Fang, F. Sun, and J. Zhang, "PointNetGPD: Detecting Grasp Configurations from Point Sets," in *International Conference on Robotics and Automation (ICRA)*, May 2019, pp. 3629–3635.

[4] P. M. Taylor, D. Pollett, and M. Grießer, "Pinching grippers for the secure handling of fabric panels," *Assembly Automation*, vol. 16, no. 3, pp. 16–21, Jan. 1996.

[5] N. Fiedler, Y. Jonetzko, and J. Zhang, "A multimodal pipeline for grasping fabrics from flat surfaces with tactile slip and fall detection," in *2023 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, 2023, pp. 1–6.

[6] F. Zhang, V. Bazarevsky, A. Vakunov, A. Tkachenka, G. Sung, C.-L. Chang, and M. Grundmann, "Mediapipe hands: On-device real-time hand tracking," *arXiv preprint arXiv:2006.10214*, 2020.