Effective grasping enables successful robot-assisted dressing

Júlia Borràs

Compared with rigid objects whose pose is fixed with position and orientation, textile objects are challenging for robots to handle because textiles change shape under contact and motion, with an infinite-dimensional configuration space. This huge dimensional jump makes existing perception and manipulation methods difficult to apply to textiles. However, these kinds of objects are very relevant in many scenarios of our daily life in domestic, health care, and industrial contexts. As a result, there has been increasing interest in the application of robots for manipulating deformable objects. The scientific community is growing; this can be seen in the number of publications, series of dedicated workshops in the past 5 years at the main robotic conferences, and the various funded projects both by the European Union and U.S. agencies. Early solutions to fold a pile of crumpled towels were very slow, and more recent ones have been successful at handling flat configurations more efficiently but with limited ability to control shape, especially under dynamic motions. Full demonstrations of assisted dressing are rare, with most demonstrations limited to partial tasks and assuming pregrasped garments. Writing in Science Robotics Zhang and Demiris [1] demonstrate how to teach a dual-arm robot to grasp hospital gowns and to dress a mannequin lying in a bed. The researchers' three key contributions are as follows: first, to improve learning from simulation examples; second, to teach grasping skills by applying pregrasp pushing motions; and third, to boost success rate to the point of showing the execution of sequences of grasps and regrasps leading to the completion of a full dressing task.

Learning how to handle textiles requires a robot to encode the way such objects deform under dynamic conditions from examples that are stored under some digital representation. These examples can only come either from cloth simulations or from real videos, and both are challenging to handle—first, because models of cloth in simulations still differ from real dynamic behavior despite their nice appearance and, more importantly, because physical parameters in available simulators are completely unrelated to real physical parameters and thus difficult to estimate. On the other hand, real data are very time-consuming to produce because robots alone cannot yet understand the cloth they see in the video, and we need to manually annotate these data. Because we need a lot of data, learning from only real examples appears to be unfeasible at the moment. Zhang and Demiris (1) show how a network can better learn from simulation data by estimating the physical parameters of cloth in an innovative way: generating many depth videos of the simulated cloth obtained with different parameters, comparing them to the video of the real one using contrastive learning, and then choosing the parameters of the simulated video that is more similar.

Pregrasp pushing motions aim to achieve natural interactions with both the object and the environment. That is a general challenge in manipulation, and to realize that, non-prehensile actions are needed to interact with the object—that is, not only grasping but also pushing and exploiting the environment to manipulate it. This is especially true for highly deformable objects as discussed in [2]. Zhang and Demiris [1] are the first to provide data examples of grasps that include pregrasp non-prehensile motions, increasing the success rate. This approach is even more relevant when compared to perception for grasping methods that just find grasping points but bypass how to approach to grasp. For instance, you may know where the correct grasping point is, even the

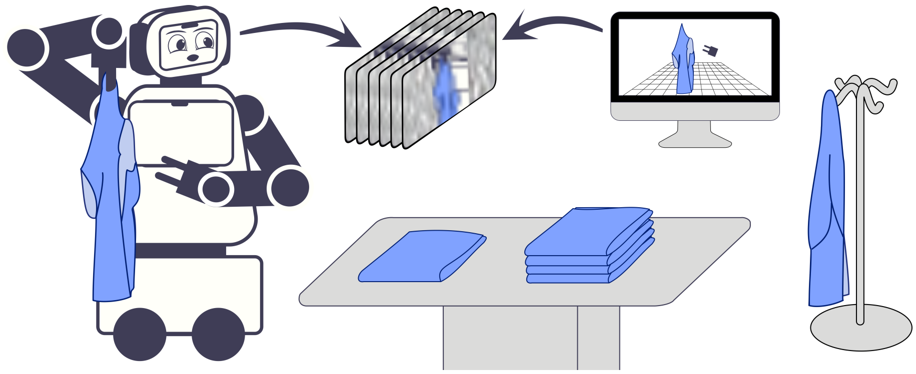

Figure 1: The robot learns to manipulate a hospital gown from simulations and real examples. Tackling the infinite dimensional configuration space of cloth is challenging because garments can be folded, flat, hanging, or grasped. Physically accurate simulations are hard to obtain because they require computation of extremely complex collisions of the textile with the environment and itself.

gripper location to grasp, but if the gripper touches the cloth before getting there, the deformation changes and so does the grasping point, leading to grasp failure. Linking shape to non-prehensile pregrasp motions and including exploitation of extrinsic constraints [3] are two key ideas underlying future perception-based systems for grasping textiles.

Another important value of the approach by Zhang and Demiris is that they are able to demonstrate the full pipeline. Robotic full pipelines are difficult to find because they require solving lots of side problems that usually are not easy to adopt because of the low repeatability in robotics. The diversity of hardware platforms and the poor generalization of many deep learning solutions are two of the main challenges roboticists will have to face in the near future. You can get a glimpse of the complexity of solving a full pipeline in table S1 of [1], showing some of the many networks that have been trained. That means lots of data collection that has to be done for each one of the objects to be grasped. We also know little about how different the three used objects are in terms of cloth properties that can affect their ability to be manipulated: rigidity, elasticity, thickness, or material. Such parameters are difficult to know from simple observation and are just starting to be considered [4, 5] but will be important for benchmarking and repeatability of results.

Many open challenges remain to have a fully applicable solution. Challenges like recovery from failures and trustworthy relationships with the patient or doctors, including both safety and trust/acceptability, are among the important ones. Facing these challenges will require control of the cloth's shape after grasping and a high-level understanding of the deformation state of the garment, an aspect that has been very unstudied. Many recent reviews, such as [6, 7], point out the need to integrate model-based and data-driven approaches and the still-open challenge to find a simplified representation of cloth that not only encodes enough from the multidimensional deformation state but also allows for a high-level understanding of such deformation in a way that can be linked to actions for decision-making. The data-driven approach of Zhang and Demiris shows impressive results, but it might be argued that some analytical model is needed to give a structure to what the robot learns, to identify similarities to reuse instead of retrain, and to be able to face unforeseen states when something goes wrong.

References

[1] F. Zhang and Y. Demiris, "Learning garment manipulation policies towards robot-assisted dressing," *Science robotics*, vol. 7, no. 65, p. abm6010, 2022.

[2] J. Borràs, G. Alenyà, and C. Torras, "A grasping-centered analysis for cloth manipulation," *IEEE Transactions on Robotics*, vol. 36, no. 3, pp. 924–936, 2020.

[3] C. Eppner, R. Deimel, J. Álvarez-Ruiz, M. Maertens, and O. Brock, "Exploitation of environmental constraints in human and robotic grasping," *International Journal of Robotic Research*, vol. 34, no. 7, pp. 1021–1038, 2015.

[4] A. Longhini, M. C. Welle, I. Mitsioni, and D. Kragic, "Textile taxonomy and classification using pulling and twisting," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 7564–7571, IEEE, 2021.

[5] I. Garcia-Camacho, J. Borràs, B. Calli, A. Norton, and G. Alenyà, "Household cloth object set: Fostering benchmarking in deformable object manipulation," *IEEE Robotics and Automation Letters (to appear)*, 2022.

[6] H. Yin, A. Varava, and D. Kragic, "Modeling, learning, perception, and control methods for deformable object manipulation," *Science Robotics*, vol. 6, no. 54, p. eabd8803, 2021.

[7] J. Sanchez, J.-A. Corrales, B.-C. Bouzgarrou, and Y. Mezouar, "Robotic manipulation and sensing of deformable objects in domestic and industrial applications: a survey," *Int. J. of Rob. Res.*, vol. 37, no. 7, pp. 688–716, 2018.