Challenges and Outlook in Robotic Manipulation of Deformable Objects

Jihong Zhu, Andrea Cherubini, Claire Dune, David Navarro-Alarcon, Farshid Alambeigi, Dmitry Berenson, Fanny Ficuciello, Kensuke Harada, Jens Kober, Xiang Li, Jia Pan, Wenzhen Yuan and Michael Gienger

Abstract—Deformable object manipulation (DOM) is an emerging research problem in robotics. The ability to manipulate deformable objects endows robots with higher autonomy and promises new applications in the industrial, services, and healthcare sectors. However, compared to rigid object manipulation, the manipulation of deformable objects is considerably more complex, and is still an open research problem. Addressing DOM challenges demand breakthroughs in almost all aspects of robotics, namely, hardware design, sensing, (deformation) modeling, planning, and control. In this article, we review recent advances and highlight the main challenges when considering deformation in each sub-field. A particular focus of our paper lies in the discussions of these challenges and proposing future directions of research.

I. INTRODUCTION

UNTIL now, object rigidity is one of the common assumptions in robotic grasping and manipulation. Strictly speaking, all objects deform upon force interaction. Rigidity is a valid assumption when object deformation can be neglected in the task. Nevertheless, many objects that need to be manipulated by robots present non-negligible deformation: from micro-surgical operation to challenging industrial assemblies.

Robots need to be capable of manipulating deformable objects to operate in human environments. This capability would benefit many application fields while posing fundamental research challenges. In this article, we consider a generalized concept of manipulation where grasping is also

J. Zhu is with Cognitive Robotics, TU Delft and Honda Research Institute, Europe. `j.zhu-3@tudelft.nl`

A. Cherubini is with LIRMM - Université de Montpellier CNRS, 161 Rue Ada, 34090 Montpellier, France. `andrea.cherubini@lirmm.fr`

C. Dune is with COSMER laboratory EA 7398, Université de Toulon, 83130 La Garde, France. `claire.dune@univ-tln.fr`

D. Navarro-Alarcon is with The Hong Kong Polytechnic University, Department of Mechanical Engineering, Kowloon, Hong Kong. `dna@ieee.org`

F. Alambeigi is with the University of Texas at Austin, Austin, USA. `farshid.alambeigi@austin.utexas.edu`

D. Berenson is with the University of Michigan, Ann Arbor, MI, USA. `dmitryb@umich.edu`

F. Ficuciello is with Università degli Studi di Napoli Federico II, 80125 Napoli, Italy `fanny.ficuciello@unina.it`

K. Harada is with the Osaka University, Japan, and the National Institute of AIST, Japan `harada@sys.es.osaka-u.ac.jp`

J. Kober is with Cognitive Robotics, TU Delft, the Netherlands `J.Kober@tudelft.nl`

X. Li is with Department of Automation, Tsinghua University, Beijing, China. `xiangli@tsinghua.edu.cn`

J. Pan is with the Department of Computer Science, The University of Hong Kong, Pok Fu Lam, Hong Kong. `jpan@cs.hku.hk`

W. Yuan is with Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213 USA `wenzheny@andrew.cmu.edu`

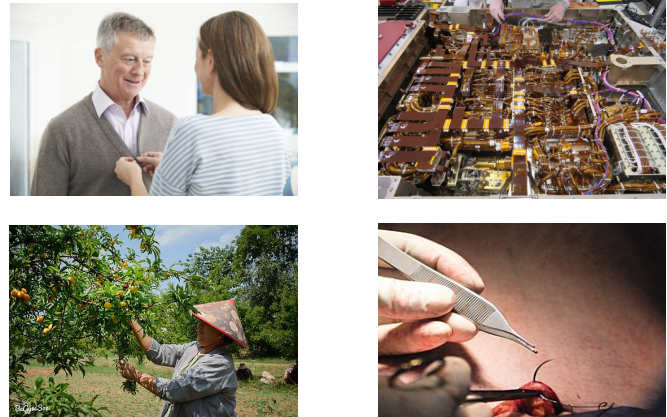M. Gienger is with Honda Research Institute Europe, Offenbach, Germany `Michael.Gienger@honda-ri.de`

Fig. 1. Applications involving manipulation of deformable objects. Clockwise from top left: dressing assistance [1], cable harnessing [2], fruit harvesting [3], suturing [4]

part of the task. We will refer to the problem as deformable object manipulation (DOM).

The tasks involved in DOM cover a broad spectrum (see Fig. 1). These include: dressing assistance in elderly care, cable harnessing in industrial automation, harvesting and processing fruit and vegetables in agriculture, surgical operations in medical services, to name a few.

On the technical side, addressing deformation introduces the following technical challenges:

- the complication of sensing deformation,
- the high number of degrees of freedom of soft bodies,
- the complexity of non-linearity in modeling deformation.

We believe that overcoming these challenges is not only beneficial to DOM, but can further push towards developing autonomous robots which can operate in unstructured environments.

In recent years, there have been a few surveys on robotic manipulation of deformable objects. Some surveys focus on specific areas of DOM. The survey from Jimenez [6] focuses on model-based manipulation planning. More recently, Herguedas et al. [7] review works using multi-robot systems for DOM while the work of [8] considers multi-modal sensing. The authors of [9] present the state-of-the-art on deformable object modeling for manipulation. There are also two comprehensive surveys in the area. The survey in [10] reviews and classifies the state-of-the-art according to the object's physical properties. Lately, [11] reports most recent advances in modeling, learning, perception, and control in DOM.
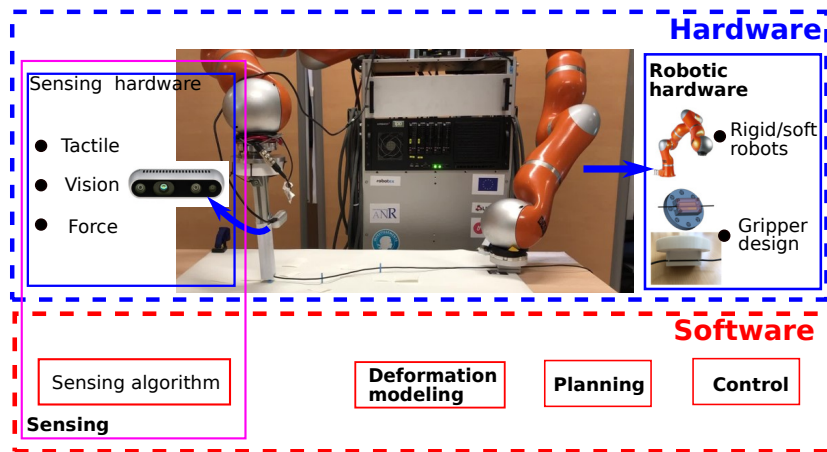
Fig. 2. A typical robotic framework for handling deformable objects. In this particular example, the framework addresses wire harness [5].

In contrast with the mentioned surveys, which either focus on reporting the progress of the field or a specific area, this article aims at identifying *scientific challenges* introduced by object deformations and at projecting crucial future research directions. As DOM is an emerging field of research where there is still much to be done, in this paper, previous works and open problems are given equal weights. In addition, we dedicate one section to discussing practical challenges in various applications of DOM. We believe the paper is a first of its kind, in the field of DOM.

A robotic framework designed to handle deformable objects usually consists of five key components: *gripper and robot design*, *sensing*, *modeling*, *planning* and *control* (Fig. 2). To position the current research and identify future trends, we conducted a survey on the future perspective of deformable object manipulation. We shared the survey with people working in related fields, at various career stages. They were asked to rate the importance and research maturity of each of the five identified key components, from 1 to 4, with 1 being not important/low maturity and 4 being very important/high maturity. We received 31 answers that are summarized on Fig. 3.

We consider the promising direction of research as the ones that have the highest significance and the lowest research maturity. Based on the survey, sensing is the most promising one among all subareas. This is probably due to the current booming trend in Deep Learning, which has offered many new methods for processing sensory data. In addition, sensing is the prerequisite for subsequent steps such as modeling, planning, and control.

Accordingly, the following sections of this paper each present one of these five research directions. In each section, we review recent works in the field and then comment on the outlook and challenges ahead. Then, Sect. VII tries to provide a link from research to practical applications in the context of DOM. Finally, we summarize key messages in Sect. VIII.
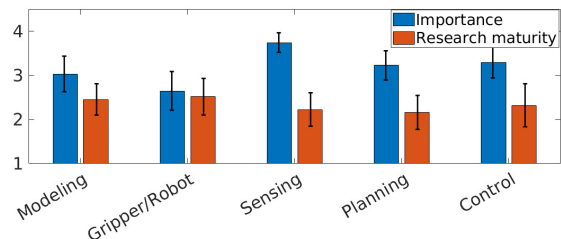
II. GRIPPER AND ROBOT DESIGN

A. Current capability

Does manipulation of deformable objects demand specific grippers as compared to manipulation of rigid objects? Gen-
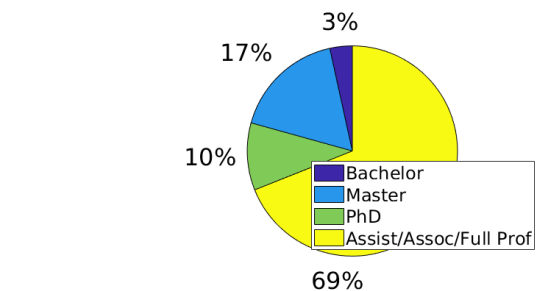
(a) Highest qualifications of the respondents.

(b) Means and variances of Importance and research maturity ratings of each key component.

Fig. 3. Summary of the outcomes of the survey on DOM. We received in total 31 answers. The respondents cover different level of qualifications ranging from master students to full professors.

erally, yes (see Fig. 4). Unlike rigid objects (which are mostly handled by standard grippers), deformable objects are handled with custom (and often designed ad-hoc) grippers, e.g. a 3D printed gripper that enables cable sliding [5], a flat clip for holding towels [12], a cylindrical tool for pushing and tapping plastic materials [13], or a soft hand for manipulating organs [14]. Such diversity in grippers is a result of the large variety of deformable objects, which require different actions during manipulation. To avoid designing task-specific grippers for DOM, human-like dexterity and compliance is desired. Recent works in this directions consider compliant design [15], [16] and show good potential for DOM tasks.

As for the robot itself, it is rigid in most works. In some cases, as in the surgical application showcased in [17] (Fig. 4, bottom right), both the robot and object are deformable to
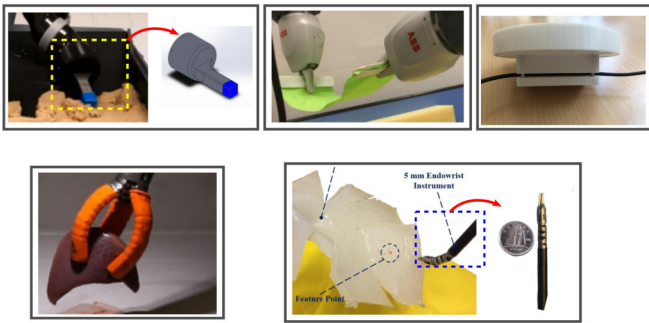
ensure the safety of manipulation.

Fig. 4. Various robot grippers for DOM. Clockwise from top left: a tool for pushing and tapping on plastic materials [13], flat clips for holding a towel [12], a gripper allowing a cable to slide [5], a soft continuum manipulator interacting with a deformable material [17] and a soft hand for manipulating organs [14].

B. Challenges and outlook

Improving dexterity is core to robot manipulation. The improvement can come from different research domains, such as accurate in-hand sensing or robust control, two aspects which we will detail in Sections III and VI, respectively. In this section, our focus is on gripper/robot hardware aspects.

One way of achieving such dexterity is to reproduce by design the most dexterous gripper – the human hand. An open question is whether the anthropomorphic design is in itself the optimal solution in all cases, especially in the context of DOM.

While having one dexterous gripper which can handle a variety of DOM tasks is appealing, it should be noted that additional constraints need to be considered in the design process, for hygiene/safety in tasks such as food handling or surgery. For instance, for surgical applications we are limited by the biocompatibility of the materials and actuators and by the reduced available space in minimally invasive surgery. In these cases, designing task-specific grippers is more appropriate. Non-anthropomorphic soft grippers are another emerging area of research [18]. These grippers are promising, to overcome the challenges associated with traditional fingered grippers in grasping rigid objects; yet, to date, their application to DOM receives little attention.

Otherwise, one may use a standard gripper, and provide the robot with suitable tools to be grasped and used according to the type of task at stake. This demands breakthroughs on the algorithmic side, to make the robot capable of reasoning about proper tools for different tasks. Training the robot to have task-specific tool reasoning will enhance autonomy and make robots realize more complex tasks.

Another area worth investigating is that of soft robots/grippers, since these have great potential for manipulating fragile materials, such as organs or food, or for collecting biological samples or fruits (see Fig. 5). While traditional rigid robots need to exhibit compliant behavior when interacting with these objects, the inherent compliance of soft robots makes the task safe. This unconventional paradigm of using soft robots to manipulate soft objects

will bring new challenges in modeling and control as both the robot and the object are under-actuated and difficult to model. One pioneering work in this direction is [21], which adapts the finite element modeling (FEM) based inverse soft robot model with contact handling (proposed in [22]) for deformable objects manipulation using soft robots.

An interesting research question to consider is whether methods can be transferred from one field to the other. To be more specific, can methods for controlling/modeling soft robots be applied to manipulating deformable objects and vice versa? If so, as a community, it may be valuable to obtain a unified approach for working with both soft robots and deformable objects.

III. SENSING

A. Current capability

In this section, we consider *visual*, *tactile* and *force* sensing for DOM. Existing research relies on these three modes to estimate the state of deformable objects. In most cases, vision provides global information about shapes on a large scale, while force and tactile provide local information on both shape and contact. At the end of this section, we also discuss the research in contrast to this common practice, where global deformation properties are recovered using tactile sensing. It should also be noted that force information is particularly important in industrial settings, e.g., for assembly [23], [24].

Vision is used in tasks such as rope manipulation [25], [26] or cloth unfolding [27], [28], where the object exhibits large global deformation. In these works, configurations of deformable objects were obtained from raw image readings. Although vision offers a global perspective of the object configuration, visual data can be noisy in unstructured environments. It is then important to manage occlusions [12], [29]. Most of the above-mentioned works are based on 2D vision, 3D perception of deformable objects is more challenging. Existing works employ FEM [30] or a combination of Growing Neural Gas and Particle Graph Networks [31] for better tracking the deformation. In a more recent study [32], it has been shown that a deep convolutional neural network for processing vision data can be used with small variations to process tactile data for deformable objects recognition. Objects made of soft materials, such as human tissues and fruits, have a special force-displacement correlation upon contact. As a result, tactile sensing can be used to estimate

(a) Picking tomatoes [19].

(b) A custom 3D printed soft robotic gripper, grasping mushroom coral [20]

Fig. 5. Two examples of interaction with fragile objects, which could benefit from the use of soft robots.

the stiffness. In [33], the GelSight [34], a vision-based high-resolution tactile sensor, measures the 3D geometry of the contact surface, and the normal/shear forces.

Note that the division of *vision for global deformation* and *tactile sensing for local deformation* is not absolute. The authors of [35] use vision to estimate the local deformation of objects during grasping, and classify objects accordingly. In [36], high-resolution tactile sensing is used to estimate the physical properties of clothing materials through squeezing, assuming the robot can learn from the data about global properties of clothing according to a local sampling point. In [37] an example of servoing along a cable based on high-resolution tactile sensing is presented. Although vision is not used, the precise measurement of the local cable shape provides enough information to guide the robot motion on a small scale.

B. Challenges and outlook

The main challenges for sensing are: selecting appropriate sensors for the DOM task and using the measurements to obtain meaningful object representations.

Considering the high number of degrees of freedom (DoF) of the deformable bodies, a fusion of different sensing modalities (vision, force and tactile) may be a promising direction to pursue in future research. Another research question to be answered is: what yields a good representation of the object configuration? We (acknowledgedly) do not have a complete answer to this; rather, we will elaborate on considerations when designing the representation.

The representation needs to be robust to noise and useful for reconstructing the objects' configuration – even when data are partially unavailable. In vision, the most common noise is occlusion. How to generate a meaningful representation of these objects under self-occlusion is still an open problem in research. For rigid objects, one can carefully design the environment to avoid it. For deformable objects that exhibit large global deformation such as clothes, bed sheets, etc, self-occlusion is inevitable during manipulation. A promising direction to deal with occlusion and noise is using active/interactive perception [38], [39]. With vision data from different perspectives, we might be able to reconstruct the object's configuration accurately even under occlusion and noise.

Apart from the above mentioned challenges, choosing a good representation also involves leveraging two aspects:

1) the dimensionality of the representation,
2) the accuracy of the representation.

Usually, the trade-off depends on the task, relies on human intuition and involves a trial and error process.

In end-to-end reinforcement learning settings, sensory data can be mapped directly to robot actions without explicit feature representations [40]. Human demonstrations can be used for making end-to-end learning more efficient. One example is reported in [41]. The authors use an improved version of Deep Deterministic Policy Gradients, trained with 20 demonstrations, to make robots manipulate cloth. However, since such settings often require a manually designed cost/reward function for learning, human demonstrations in this context

can also be used for recovering the reward with inverse reinforcement learning.

IV. MODELING

A. Current capability

For robots to perform deformation tasks using sensory data, we need a model that captures the relationship between sensor information and robot motion. A linear model characterized by Young's modulus can be employed for describing elastic deformation. The two other classes of deformation are plastic, and elasto-plastic deformations. This classification serves well. However, since the model should be used for control, in this section, we prefer to distinguish between local and global models – a taxonomy that has clearer implications for control. We introduce the corresponding research and – at the end of the section – we discuss the limitations of these models and present works that address them.

Most local models approximate the perception/action relationship via a Jacobian Matrix (called *Interaction Matrix* in visual servoing). Such a model is linear and can be computed in real-time with a small amount of data. However, since it is a local model, it must be continuously updated during task execution. Model updating methods include: Broyden rule [17], receding horizon adaption [42], local gradient descent [43], QP-based optimization [44], and Multi-armed Bandit-based methods [45]. Another advantage of the Jacobian model is that one can design a simple controller by inverting it. However, since this controller is local, it operates via a series of intermediate target shapes [42], [44].

On the other hand, global models can be approximated with Finite Element Methods [46] and also (deep) neural networks. In contrast to simple linear models, (D)NN-based approaches benefit from stronger representation power, in terms of accuracy and robustness [47]. Moreover, they can incorporate physics models and reason about object interaction [48]. These models can approximate highly nonlinear systems and have a larger validity range, solving (to some extent) the locality issue of the linear models. Nevertheless, these complex nonlinear representations demand large amounts of data (which might not be available in some cases).

Whether analytical or learned models are used, their predictive power will be limited. They are either specialized to some class of tasks or learned from a set of training data. Especially for the learned models, we can never hope to collect enough data to produce an accurate model in the entire state space (which is high dimensional). Thus [49] and [50] have developed methods to reason about the validity of a (learned) model for a given state and action, and have used these methods to reason about model uncertainty in planning and control. However, when the model is not precise, a re-planning/recovery might be desirable. The authors of [51] introduce two neural networks for learning and re-planning the motion when the model is unreliable.

B. Challenges and outlook

The complexity of modeling is manifested in the lack of simulators. While most existing robotic simulators are capable of producing rigid body kinematics and dynamic behaviours,

only a fraction of them can handle deformation. One recent work, Softgym [52] was proposed for bench-marking DOM based on Nvidia Flex. In the soft robotics community, SOFA [53] and Chainqueen [54] are example simulators. In Sect. II-B, we considered the interaction between soft robots and deformable objects. Thus, a unified simulator that is able to handle soft robots and objects, and model their interaction might be desirable.

When choosing a model for control, one challenge of data-driven deformation modeling is to balance the region of validity with the amount of data required for training. One possible direction is to combine a simple model with a complex nonlinear model to form a hierarchical model. An example of such structures is exploited in [55] for robust in-hand manipulation. For DOM tasks, we can have a linear model at the lower level, and a (D)NN learning the full model at a higher level. The lower level model can be learned in few iterations to enable instant interaction between the robot and object. The higher level (D)NN can collect data and improve the model to enhance global convergence.

V. PLANNING

A. Current capability

Planning aims at finding a sequence of valid (robot/object) configurations and contributes to solving the problem of limited validity of local models, as discussed in Sect. IV.

Planners can operate in the objects' configuration space, and sometimes rely heavily on physics-based simulation. While the obtained plan can be visually plausible, it may be unrealizable for a specific object. Recently, McConachie et al. presented a framework that combines global planning without physics simulation, with local control [56]. For an elastic object, considering its energy is another way to do planning; in this direction, Ramirez-Alpizar et al. [57] proposed a dual-arm manipulation planner optimizing the elastic energy, for elastic ring-shaped objects manipulation. For DOM tasks involving multiple robots, planning is important for coordination. Alonso-Mora et al. employed a distributed receding horizon planner for transporting tasks that require multiple robots [58]. More recently, [59] learns a latent representation for semantic soft object manipulation that enables (quasi) shape planning with deformable objects.

With Learning from Demonstration (LfD), the robot can be trained to manipulate deformable objects by an expert (usually a human). LfD encodes the robot trajectory and interaction force from human demonstrations [60], thus avoiding explicitly planning the motion. More recently, Wu et al. have proposed a reinforcement learning scheme for DOM, which does not require initial demonstrations [61].

B. Challenges and outlook

A rigid object configuration can be described in space with 6 DoF, whereas a deformable object configuration has a much higher number of DoF. To address this from the sensing algorithm side, one can find a compact representation from sensory data, as discussed in Sect. III-B. An alternative, which receives much less attention, is the use of environmental contacts to constrain some DoF of deformable objects. Examples include the use of contact points in cable harnesses or that of flat surfaces when folding clothes. We argue that instead of planning to avoid contacts as most planners do, for deformable objects, we need to plan to make contact, since this constrains the configuration, and therefore simplifies the task.

Planning to grasp the correct point is often crucial in DOM tasks. For instance, grasping at convex vertices of the clothes guarantees stability and facilitates the task [62]. Re-grasp planning is highly relevant when considering tasks that require multiple robotic arms. Additional challenges come from perception, since as soon as the robot releases one or more grasp(s), the object is likely to change its configuration. We rely on sensing to track configuration changes and then plan accordingly.

Another important future work in planning is reasoning about a deformable object at a semantic level. What does it mean for a cloth to be *folded*? What does it mean for an object to be *wrapped* in a paper? We cannot manually specify all the configurations of the deformable object to use as goals in these kinds of tasks. Instead, we need a way to learn the meaning of semantic concepts, such as *folded* or *wrapped*, so that we can determine if a given configuration of the object is a valid goal.

VI. CONTROL

A. Current capability

Control aims at designing inputs for the robot to realize the planned motion. The type of controller is decided usually by the task. For instance, the authors employed a data-driven model predictive control [63] for cutting considering its predictive nature and the lower demand for manual tuning. For safe interaction in minimally invasive surgery, the authors of [64] used a fuzzy compensator with impedance control. For controlling large deformation, Aranda et al. proposed a Shape-from-Template algorithm concerning its low dimensional representation (using the template) and robustness against occlusion [65].

Several works focus on shape control. While global models directly map sensor data to robot motion, local models must be inverted to design the robot motion controller (see Sec. IV). Several applications of the control scheme for robotic manipulation of deformable objects can be found in 3C manufacturing [66], [67], where vision-based controllers were proposed to drive the robot to automatically grasp/contact the deformable object, then carry out the task of active deformation or separation/sorting. Other works consider the concept of diminishing rigidity to do deformation control [68], [69].

B. Challenges and outlook

Feedback control has been commonly used in most DOM works, by referring to the state of the object, to achieve the task. Note that such state is retrieved from the output of its deformation model and measured with sensors, and that output and state do not necessarily have the same representation and dimension. Furthermore, we can distinguish between model-based and model-free control. Due to the complexity of modeling the deformation, when using the model to derive

control policies, the controller has to take into account that the model will be inaccurate or even wrong.

Model-free approaches do not require information about the deformation parameters or the structure of the deformation model. Examples include LfD or (Deep) reinforcement learning, where the challenges are: efficient use of data, and policy generalization. To address these issues, we can combine offline and online learning methods. In the offline phase, the supervised network can be trained to estimate the model, by collecting pairs of a series of predefined inputs (e.g., the velocity of the robot end-effector) and the deformation of the object. The estimated model in the offline phase can be further updated online during the control task with adaption techniques (e.g., the adaptive NNs), to compensate for the errors due to insufficient training in the offline phase or the changes of the deformation model. Hence, both complement each other.

When multiple features on the deformable object are controlled in parallel, the system becomes under-actuated, with fewer control inputs than error outputs. Then, the robot controller should be able to deal with the conflicts between multiple features or decouple the control of multiple features in a sequential manner, to guarantee controllability.

In addition, due to the deformation during control, the contact between robot end-effectors and deformable objects may not always be maintained. Most existing systems require a certain level of human assistance to initiate the contact or to re-establish it, if it is lost during the task. To improve autonomy, the robot controller should automatically grasp or touch the object first, whenever physical contact is lost, laying the foundation of the subsequent manipulation task. Such a capability would allow the robot to effectively deal with the unforeseen changes due to deformation.

VII. PRACTICAL APPLICATIONS

In the previous sections, we centered our discussion from a scientific point of view, here, we instead discuss challenges in various applications of DOM.

Automatic laundry: A typical domestic application of DOM is laundry folding. A Tokyo-based company unveiled its prototype laundry-folding robot in 2015. However, the company was announced bankrupt in 2019 due to lack of funding for development and difficulties in improving the robot to reach a satisfactory level [70]. Although cloth folding has been tackled in a few previous research [71]–[74], it remains largely a laboratory product (limited to structured environments, certain types of clothes, etc). Commercializing the technology seems to require substantial efforts.

Assistive dressing: Robotic dressing assistance has the potential to become an important technology due to the pressing need for ageing society support. Research can roughly be categorized into simulation-based learning [75], [76] and imitation learning [77] approaches. Examples are dressing support for shoes [78], shirts [79]–[81] and pants (An example of shirt dressing is shown in Fig. 6a). However, several technical and societal challenges have to be addressed before robot-assisted dressing will become a broadly used DOM technology: physical safety for the human, modeling and prediction of the human-robot interaction, robustness for large
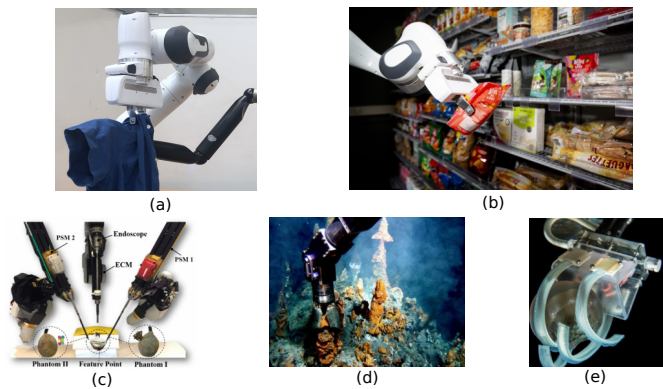

Fig. 6. Various applications of DOM – (a): A mock-up for robotics dressing assistance, (b): a robot picking a flexible bag of goods on the shelf, (courtesy AIRLab TU Delft, photo by Guus Schoonewille), (c): autonomous surgical manipulation by the dVRK system [83], (d): ROV Victor 6000 sampling black smokers (IFREMER/GENAVIR) courtesy D. Desbruyères, (e): Ultra soft underwater gripper for jellyfish [84]

variations of geometric and dynamic properties of textiles, low-cost high-reliable robot hardware, and human acceptance of such technologies.

Surgical robotics: Soft tissue manipulation is mainly performed with teleoperation solely using visual feedback. Autonomous manipulation, however, still has a long way to go and demands developing various DOM hardware and software (Fig. 6c). The biggest concern for an autonomous solution is the safety of operation. A soft robot with intrinsic compliance will probably enhance safety.

Food production & Retail: Handling deformable objects is a major challenge in the whole chain from production to sales. In an agricultural setting, automated harvesting of fruits and vegetables requires interactions with deformable objects that are at the same time easy to damage, which immediately decreases their value and shelf live. Frequently, these products also undergo an intermediate processing step (e.g., filleting and packaging meat). More generally, deformable products (e.g., everything packaged in flexible bags, (Fig. 6b)) need to be handled in warehouses, in order picking, and in restocking. Solutions for specific applications and products have been developed, but more complex objects and operations still are frequently handled by human workers.

Marine robotics: Underwater grasping has been led by the oil and gas industry for decades, resulting in heavy machines with strong grippers for inspection and maintenance tasks (Fig. 6d). Gradually the demands turned to more detailed tasks in marine biology, sedimentology, and archaeology (Fig. 6e). Another DOM application can be found in tethered robot umbilical modeling and control. Negative buoyancy cable can be modeled in real-time as a simple catenary shape and tracked to control a tethered ROV [82].

VIII. SUMMARY AND KEY MESSAGES

The revolution of robots from automating repetitive tasks to humanizing robot behaviours is taking place with better hardware, robust sensing capabilities, accurate modeling, increasingly versatile planning, and advanced control. Manipulation of deformable objects breaks fundamental assumptions

in robotics such as rigidity, known dynamics models and low-dimensional state space. It, therefore, requires breakthroughs in all the areas mentioned above, and serves as a great test-bench for novel ideas in both robotic hardware and software.

In terms of hardware, recently, the community has been shifting more and more from rigid to soft robots. Robotic manipulation is also gradually shifting from rigid to deformable objects. One open question is if some of the algorithms in one field are transferable to the other? We believe the interaction between a soft robot and a deformable object will bring more challenges to the robotic community.

Sensing plays a vital part in the robotics manipulation of deformable objects. Depending on the nature and complexity of the task, one or multiple fused sensing modes may be needed. Machine learning will facilitate the development of robust algorithms to process data from different sensors, to generate meaningful representations of deformation.

All models are wrong, some are useful. We do not believe there exists the "best" model for deformation. While more and more models tend to be data-driven, we would like to draw the readers' attention to the importance of physical models for studying interactions.

For planning, current research lacks high-level semantic reasoning of the DOM task. Furthermore, while often the purpose of planning is to avoid contact and collision, we argue that for DOM, it can be very useful to plan for contact.

Under-actuation is a key challenge of DOM, due to the deformable bodies' high DoF. Another practical issue introduced with deformation is contact loss during manipulation; future controllers should be able to detect contact loss and react accordingly.

REFERENCES

[1] Highwaystarz Photography, "Adult daughter helping senior man to button cardigan." [Online]. Available: https://www.istockphoto.com/en/photo/adult-daughter-helping-senior-man-to-button-cardigan-gm517973558-89760669

[2] NASA/JPL-Caltech, "Mars 2020 rover–inverted body wiring," June 2019. [Online]. Available: https://commons.wikimedia.org/wiki/File:PIA23312-Mars2020-Rover-InvertedBodyWiring-20190601.jpg

[3] Ba Gyee Soe, "A woman harvesting seasonal fruit in a garden," May 2020. [Online]. Available: https://commons.wikimedia.org/wiki/File:A_woman_harvesting_seasonal_fruit_in_a_garden_(May_2020).jpg

[4] James C. Mutter, "Plastic surgeon Vishal Kapoor, MD performing abdominoplasty surgery." Oct 2010. [Online]. Available: https://commons.wikimedia.org/wiki/File:Vishal_Kapoor_MD_TummyTuck_Suture.jpg

[5] J. Zhu, B. Navarro, R. Passama, P. Fraisse, A. Crosnier, and A. Cherubini, "Robotic manipulation planning for shaping deformable linear objects with environmental contacts," *IEEE Robotics and Automation Letters*, vol. 5, no. 1, pp. 16–23, 2019.

[6] P. Jiménez, "Survey on model-based manipulation planning of deformable objects," *Robotics and Computer-integrated Manufacturing*, vol. 28, no. 2, pp. 154–163, 2012.

[7] R. Herguedas, G. López-Nicolás, R. Aragüés, and C. Sagüés, "Survey on multi-robot manipulation of deformable objects," in *2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*. IEEE, 2019, pp. 977–984.

[8] F. Nadon, A. J. Valencia, and P. Payeur, "Multi-modal sensing and robotic manipulation of non-rigid objects: A survey," *Robotics*, vol. 7, no. 4, p. 74, 2018.

[9] V. E. Arriola-Rios, P. Guler, F. Ficuciello, D. Kragic, B. Siciliano, and J. L. Wyatt, "Modeling of deformable objects for robotic manipulation: A tutorial and review," *Frontiers in Robotics and AI*, vol. 7, 2020.

[10] 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 Robotics Research*, 2018.

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

[12] 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*, vol. 4, no. 4, pp. 4255–4261, 2019.

[13] A. Cherubini, V. Ortenzi, A. Cosgun, R. Lee, and P. Corke, "Model-free vision-based shaping of deformable plastic materials," *The Int. J. of Robotics Research*, 2020.

[14] H. Liu, M. Selvaggio, P. Ferrentino, R. Moccia, S. Pirozzi, U. Bracale, and F. Ficuciello, "The musha hand ii: A multi-functional hand for robot-assisted laparoscopic surgery," *IEEE/ASME Transactions on Mechatronics*, 2020.

[15] C. Della Santina, G. Grioli, M. Catalano, A. Brando, and A. Bicchi, "Dexterity augmentation on a synergistic hand: the pisa/iit softhand+," in *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*. IEEE, 2015, pp. 497–503.

[16] S. Abondance, C. B. Teeple, and R. J. Wood, "A dexterous soft robotic hand for delicate in-hand manipulation," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5502–5509, 2020.

[17] F. Alambeigi, Z. Wang, R. Hegeman, Y.-H. Liu, and M. Armand, "Autonomous data-driven manipulation of unknown anisotropic deformable tissues using unmodelled continuum manipulators," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 254–261, 2018.

[18] Y. Hao and Y. Visell, "Beyond soft hands: Efficient grasping with non-anthropomorphic soft grippers," *Frontiers in Robotics and AI*, vol. 8, p. 210, 2021. [Online]. Available: https://www.frontiersin.org/article/10.3389/frobt.2021.632006

[19] Kinwun, "Image processing technology was applying with the robot used to harvest tomatoes in the agricultural sector." [Online]. Available: https://www.istockphoto.com/en/photo/image-processing-technology-was-apply-with-the-robot-to-used-to-harvesting-tomatoes-gm1130498724-299005093

[20] D. M. Vogt, K. P. Becker, B. T. Phillips, M. A. Graule, R. D. Rotjan, T. M. Shank, E. E. Cordes, R. J. Wood, and D. F. Gruber, "Shipboard design and fabrication of custom 3d-printed soft robotic manipulators for the investigation of delicate deep-sea organisms," *PLOS ONE*, vol. 13, no. 8, pp. 1–16, 08 2018.

[21] F. Ficuciello, A. Migliozzi, E. Coevoet, A. Petit, and C. Duriez, "Fem-based deformation control for dexterous manipulation of 3d soft objects," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 4007–4013.

[22] E. Coevoet, A. Escande, and C. Duriez, "Optimization-based inverse model of soft robots with contact handling," *IEEE Robotics and Automation Letters*, vol. 2, no. 3, pp. 1413–1419, 2017.

[23] J. Luo, E. Solowjow, C. Wen, J. A. Ojea, and A. M. Agogino, "Deep reinforcement learning for robotic assembly of mixed deformable and rigid objects," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 2062–2069.

[24] Y. Hayami, P. Shi, W. Wan, I. G. Ramirez-Alpizar, and K. Harada, "Multi-dimensional error identification during robotic snap assembly," in *IFToMM World Congress on Mechanism and Machine Science*. Springer, 2019, pp. 2189–2198.

[25] A. Nair, D. Chen, P. Agrawal, P. Isola, P. Abbeel, J. Malik, and S. Levine, "Combining self-supervised learning and imitation for vision-based rope manipulation," in *2017 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2017, pp. 2146–2153.

[26] M. Yan, Y. Zhu, N. Jin, and J. Bohg, "Self-supervised learning of state estimation for manipulating deformable linear objects," *IEEE robotics and automation letters*, vol. 5, no. 2, pp. 2372–2379, 2020.

[27] L. Sun, G. Aragon-Camarasa, S. Rogers, and J. P. Siebert, "Accurate garment surface analysis using an active stereo robot head with application to dual-arm flattening," in *2015 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2015, pp. 185–192.

[28] Y. Li, D. Xu, Y. Yue, Y. Wang, S.-F. Chang, E. Grinspun, and P. K. Allen, "Regrasping and unfolding of garments using predictive thin shell modeling," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2015, pp. 1382–1388.

[29] C. Cheng and D. Berenson, "Occlusion-robust deformable object tracking without physics simulation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019.

[30] A. Petit, V. Lippiello, G. A. Fontanelli, and B. Siciliano, "Tracking elastic deformable objects with an rgb-d sensor for a pizza chef robot," *Robotics and Autonomous Systems*, vol. 88, pp. 187–201, 2017.

[31] A. J. Valencia, F. Nadon, and P. Payeur, "Toward real-time 3d shape tracking of deformable objects for robotic manipulation and shape control," in *2019 IEEE SENSORS*. IEEE, 2019, pp. 1–4.

[32] G. Rouhafzay, A.-M. Cretu, and P. Payeur, "Transfer of learning from vision to touch: A hybrid deep convolutional neural network for visuo-tactile 3d object recognition," *Sensors*, vol. 21, no. 1, p. 113, 2021.

[33] W. Yuan, M. A. Srinivasan, and E. H. Adelson, "Estimating object hardness with a gelsight touch sensor," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2016, pp. 208–215.

[34] W. Yuan, S. Dong, and E. H. Adelson, "Gelsight: High-resolution robot tactile sensors for estimating geometry and force," *Sensors*, vol. 17, no. 12, p. 2762, 2017.

[35] P. Güler, Y. Bekiroglu, X. Gratal, K. Pauwels, and D. Kragic, "What's in the container? classifying object contents from vision and touch," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2014, pp. 3961–3968.

[36] W. Yuan, Y. Mo, S. Wang, and E. H. Adelson, "Active clothing material perception using tactile sensing and deep learning," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 4842–4849.

[37] Y. She, S. Wang, S. Dong, N. Sunil, A. Rodriguez, and E. Adelson, "Cable manipulation with a tactile-reactive gripper," in *Robotics: Science and Systems*, 2020.

[38] J. Bohg, K. Hausman, B. Sankaran, O. Brock, D. Kragic, S. Schaal, and G. S. Sukhatme, "Interactive perception: Leveraging action in perception and perception in action," *IEEE Transactions on Robotics*, vol. 33, no. 6, pp. 1273–1291, 2017.

[39] R. Bajcsy, Y. Aloimonos, and J. K. Tsotsos, "Revisiting active perception," *Autonomous Robots*, vol. 42, no. 2, pp. 177–196, 2018.

[40] S. Levine, C. Finn, T. Darrell, and P. Abbeel, "End-to-end training of deep visuomotor policies," *J. of Machine Learning Research*, vol. 17, no. 39, pp. 1–40, 2016.

[41] J. Matas, S. James, and A. J. Davison, "Sim-to-real reinforcement learning for deformable object manipulation," in *Conference on Robot Learning*. PMLR, 2018, pp. 734–743.

[42] J. Zhu, D. Navarro-Alarcon, R. Passama, and A. Cherubini, "Vision-based manipulation of deformable and rigid objects using subspace projections of 2d contours," *Robotics and Autonomous Systems*, vol. 142, 2021.

[43] D. Navarro-Alarcon and Y.-H. Liu, "Fourier-based shape servoing: A new feedback method to actively deform soft objects into desired 2-D image contours," *IEEE Trans. on Robotics*, vol. 34, no. 1, pp. 272–1279, 2018.

[44] R. Lagneau, A. Krupa, and M. Marchal, "Automatic shape control of deformable wires based on model-free visual servoing," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5252–5259, 2020.

[45] D. McConachie and D. Berenson, "Bandit-Based Model Selection for Deformable Object Manipulation," *Workshop on the Algorithmic Foundations of Robotics (WAFR)*, 2016.

[46] E. Yoshida, K. Ayusawa, I. G. Ramirez-Alpizar, K. Harada, and C. Duriez, "Simulation-based optimal motion planning for deformable object," in *2015 IEEE International Workshop in Advanced Robotics and its Social Impacts (ARSO)*, 2015.

[47] A. J. Valencia and P. Payeur, "Combining self-organizing and graph neural networks for modeling deformable objects in robotic manipulation," *Frontiers in Robotics and AI*, vol. 7, 2020.

[48] P. W. Battaglia, R. Pascanu, M. Lai, D. Rezende, and K. Kavukcuoglu, "Interaction networks for learning about objects, relations and physics," 2016.

[49] D. McConachie, T. Power, P. Mitrano, and D. Berenson, "Learning When to Trust a Dynamics Model for Planning in Reduced State Spaces," *IEEE Robotics and Automation Letters (RA-L)*, 2020.

[50] T. Power and D. Berenson, "Keep it Simple: Data-efficient Learning for Controlling Complex Systems with Simple Models," *IEEE Robotics and Automation Letters (RA-L)*, 2021.

[51] P. Mitrano, D. McConachie, and D. Berenson, "Learning where to trust unreliable models in an unstructured world for deformable object manipulation," *Science Robotics*, vol. 6, no. 54, p. eabd8170, 2021.

[52] X. Lin, Y. Wang, J. Olkin, and D. Held, "Softgym: Benchmarking deep reinforcement learning for deformable object manipulation," *arXiv preprint arXiv:2011.07215*, 2020.

[53] E. Coevoet, T. Morales-Bieze, F. Largilliere, Z. Zhang, M. Thieffry, M. Sanz-Lopez, B. Carrez, D. Marchal, O. Goury, J. Dequidt *et al.*, "Software toolkit for modeling, simulation, and control of soft robots," *Advanced Robotics*, vol. 31, no. 22, pp. 1208–1224, 2017.

[54] Y. Hu, J. Liu, A. Spielberg, J. B. Tenenbaum, W. T. Freeman, J. Wu, D. Rus, and W. Matusik, "Chainqueen: A real-time differentiable physical simulator for soft robotics," in *2019 International conference on robotics and automation (ICRA)*. IEEE, 2019, pp. 6265–6271.

[55] T. Li, K. Srinivasan, M. Q.-H. Meng, W. Yuan, and J. Bohg, "Learning hierarchical control for robust in-hand manipulation," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 8855–8862.

[56] D. McConachie, A. Dobson, M. Ruan, and D. Berenson, "Manipulating deformable objects by interleaving prediction, planning, and control," *arXiv preprint arXiv:2001.09950*, 2020.

[57] I. Ramirez-Alpizar, K. Harada, and E. Yoshida, "Motion planning for dual-arm assembly of ring-shaped elastic objects," in *IEEE-RAS International Conference on Humanoid Robots*, 2014, pp. 594–600.

[58] J. Alonso-Mora, R. Knepper, R. Siegwart, and D. Rus, "Local motion planning for collaborative multi-robot manipulation of deformable objects," in *2015 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2015, pp. 5495–5502.

[59] P. Zhou, J. Zhu, S. Huo, and D. Navarro-Alarcon, "LaSeSOM: A latent and semantic representation framework for soft object manipulation," *IEEE Robotics and Automation Letters*, 2021.

[60] A. X. Lee, H. Lu, A. Gupta, S. Levine, and P. Abbeel, "Learning force-based manipulation of deformable objects from multiple demonstrations," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2015, pp. 177–184.

[61] Y. Wu, W. Yan, T. Kurutach, L. Pinto, and P. Abbeel, "Learning to manipulate deformable objects without demonstrations," *arXiv preprint arXiv:1910.13439*, 2019.

[62] J. Van Den Berg, S. Miller, K. Goldberg, and P. Abbeel, "Gravity-based robotic cloth folding," in *Algorithmic Foundations of Robotics IX*. Springer, 2010, pp. 409–424.

[63] I. Mitsioni, Y. Karayiannidis, J. A. Stork, and D. Kragic, "Data-driven model predictive control for the contact-rich task of food cutting," in *2019 IEEE-RAS 19th International Conference on Humanoid Robots (Humanoids)*, 2019, pp. 244–250.

[64] H. Su, C. Yang, G. Ferrigno, and E. De Momi, "Improved human–robot collaborative control of redundant robot for teleoperated minimally invasive surgery," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1447–1453, 2019.

[65] M. Aranda, J. A. C. Ramon, Y. Mezouar, A. Bartoli, and E. Özgür, "Monocular visual shape tracking and servoing for isometrically deforming objects," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 7542–7549.

[66] X. Li, X. Su, and Y.-H. Liu, "Vision-based robotic manipulation of flexible pcbs," *IEEE/ASME Transactions on Mechatronics*, vol. 23, no. 6, pp. 2739–2749, 2018.

[67] X. Li, X. Su, Y. Gao, and Y.-H. Liu, "Vision-based robotic grasping and manipulation of usb wires," in *IEEE International Conference on Robotics and Automation*, 2018, pp. 3482–3487.

[68] F. Nadon and P. Payeur, "Grasp selection for in-hand robotic manipulation of non-rigid objects with shape control," in *2020 IEEE International Systems Conference (SysCon)*. IEEE, 2020, pp. 1–8.

[69] M. Ruan, D. Mc Conachie, and D. Berenson, "Accounting for directional rigidity and constraints in control for manipulation of deformable objects without physical simulation," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 512–519.

[70] K. Nagata, "Laundry-phobics' dreams crushed as tokyo-based developer of laundroid robot files for bankruptcy," Apr 2019. [Online]. Available: https://www.japantimes.co.jp/news/2019/04/24/business/corporate-business/tokyo-based-developer-laundry-folding-robot-files-bankruptcy/

[71] A. Verleysen, M. Biondina, and F. Wyffels, "Video dataset of human demonstrations of folding clothing for robotic folding," *The International Journal of Robotics Research*, vol. 39, no. 9, pp. 1031–1036, 2020.

[72] S. Miller, J. Van Den Berg, M. Fritz, T. Darrell, K. Goldberg, and P. Abbeel, "A geometric approach to robotic laundry folding," *Int. J. of Robotics Research*, vol. 31, no. 2, pp. 249–267, 2012.

[73] Y. Tsurumine, Y. Cui, E. Uchibe, and T. Matsubara, "Deep reinforcement learning with smooth policy update: Application to robotic cloth manipulation," *Robotics and Autonomous Systems*, vol. 112, pp. 72–83, 2019.

[74] 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.

[75] A. Clegg, W. Yu, J. Tan, C. K. Liu, and G. Turk, "Learning to dress: Synthesizing human dressing motion via deep reinforcement learning," *ACM Transactions on Graphics (TOG)*, vol. 37, no. 6, pp. 1–10, 2018.

[76] A. Clegg, Z. Erickson, P. Grady, G. Turk, C. C. Kemp, and C. K. Liu, "Learning to collaborate from simulation for robot-assisted dressing," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2746–2753, 2020.

[77] R. P. Joshi, N. Koganti, and T. Shibata, "A framework for robotic clothing assistance by imitation learning," *Advanced Robotics*, vol. 33, no. 22, pp. 1156–1174, 2019.

[78] G. Canal, G. Alenyà, and C. Torras, "Adapting robot task planning to user preferences: an assistive shoe dressing example," *Autonomous Robots*, vol. 43, no. 6, pp. 1343–1356, 2019.

[79] S. Li, N. Figueroa, A. Shah, and J. A. Shah, "Provably Safe and Efficient Motion Planning with Uncertain Human Dynamics," in *Proceedings of Robotics: Science and Systems*, Virtual, July 2021.

[80] T. Tamei, T. Matsubara, A. Rai, and T. Shibata, "Reinforcement learning of clothing assistance with a dual-arm robot," in *2011 11th IEEE-RAS International Conference on Humanoid Robots*. IEEE, 2011, pp. 733–738.

[81] F. Zhang, A. Cully, and Y. Demiris, "Probabilistic real-time user posture tracking for personalized robot-assisted dressing," *IEEE Transactions on Robotics*, vol. 35, no. 4, pp. 873–888, 2019.

[82] M. Laranjeira, C. Dune, and V. Hugel, "Catenary-based visual servoing for tether shape control between underwater vehicles," *Ocean Engineering*, vol. 200, p. 107018, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0029801820300949

[83] F. Alambeigi, Z. Wang, R. Hegeman, Y.-H. Liu, and M. Armand, "A robust data-driven approach for online learning and manipulation of unmodeled 3-d heterogeneous compliant objects," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 4140–4147, 2018.

[84] N. R. Sinatra, C. B. Teeple, D. M. Vogt, K. K. Parker, D. F. Gruber, and R. J. Wood, "Ultragentle manipulation of delicate structures using a soft robotic gripper," *Science Robotics*, vol. 4, no. 33, 2019.