Challenges and Outlook in Robotic Manipulation of Deformable Objects

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

2022

Preprint:

This is an accepted article published in Robotics and Automation Magazine (RAM). The final authenticated version is available online at: https://doi.org/10.1109/MRA.2022.3147415

Challenges and Outlook in Robotic Manipulation of Deformable Objects

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

***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-neglectable deformation: from micro surgical operation to challenging industrial assemblies or marine nature sample collection.

While you are looking at this paper on your computer, several deformable objects are within your sight — the power cable that charges your computer, fabrics that make your cloths, papers on the table. 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 part of the task. We will refer to the problem as deformable object manipulation (DOM).

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

The tasks involved in DOM cover a broad spectrum (with examples shown in 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

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, Universié 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. berenson@eecs.umich.edu

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

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

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 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

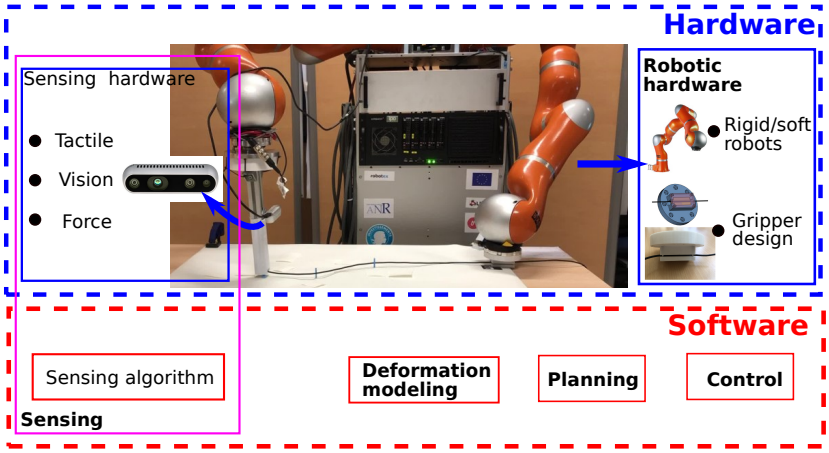


Fig. 2. A typical robotic framework for handling deformable objects. In this particular example, the framework addresses wire harness [5].

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.

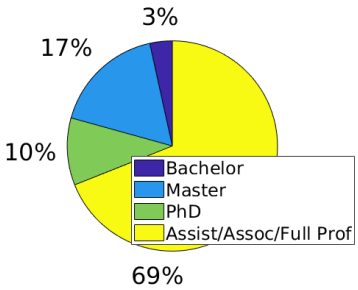
In contrast with the mentioned surveys, which either focus on reporting the progress of the field or a specific area, this article aims at identifying *scientific challenges* introduced by object deformations and at projecting crucial future research directions. As DOM is an emerging field of research 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: *hardware*, *sensing*, *modeling*, *planning* and *control* (See Fig. 2). To position the current research and identify future trends, we conducted a survey on the future perspective of deformable object manipulation[1]. We shared the survey with people working in related field, at various carrier stage (ranging from master student to full professor). They were asked to rate the importance and research maturity of each of the five idendified 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.

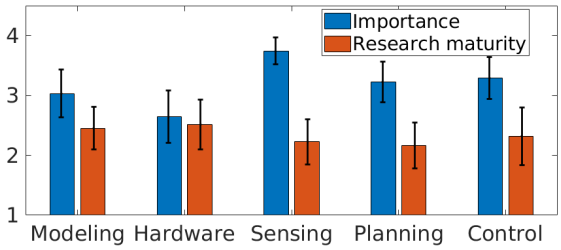
We consider the promising direction of research as the ones that have the highest significance and the lowest research maturity. Based on the survey, while all subareas are of relevance and require further investigation, sensing is undoubtedly the most promising one. This is probably due to current booming trend in Deep Learning has offered enormous new methods for processing the 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 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.

[1]Link to the survey: https://forms.gle/XCv2CV79yvRP5Gsd7



(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.

II. Hardware

A. Current capability

Does manipulation of deformable objects demand specific grippers as compared to that of rigid objects? Generally, 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], a soft hand for manipulating organs [14] or bilogical undewater sampleos [15].

As for the robot itself, it is rigid in most works. In some cases, as in the surgical application showcased in [16] (Fig. 4, bottom right), both robot and object are deformable to ensure safety of manipulation.

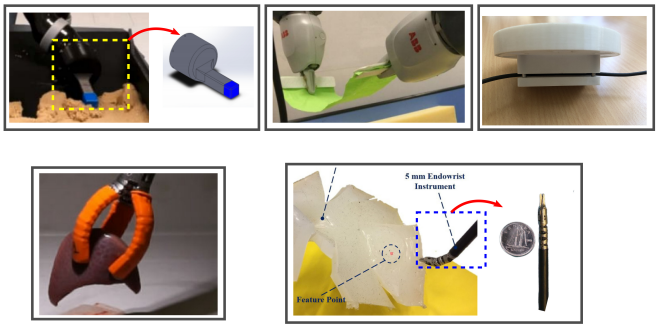


Fig. 4. Various robot hardware for DOM. Clockwise from top left: a tool for pushing and tapping on plastic materials [13], flat clips for holding a towel [12], a gripper allowing a cable to slide [5], a soft hand for manipulating organs [14], and a soft continuum manipulator interacting with a deformable material [16].

B. Challenges and outlook

Improving dexterity is core to manipulation research. The improvement can come from different research domains such as accurate in-hand sensing (discussed in Sect. III), robust control (discussed in Sect. VI) or better gripper hardware. In this section, our focus is on the hardware design aspect.

The community demands novel grippers, capable of interacting with diverse deformable objects. Search algorithms can be used to search and aid the design of the gripper.

Instead of relying on a versatile gripper, one may use a standard gripper and provide suitable tools for manipulation. This demands breakthroughs on the algorithmic side, to make the robot reasoning on the proper tools for different tasks. Training the robot to have task specific reasoning will enhance autonomy and make robots realize more complex tasks.

Another area worth investigating is that of soft robots, 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.

(a) Picking raspberries [17].

(b) A custom 3D printed soft robotic gripper, grasping mushroom coral [15]

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

Another interesting research question to consider is whether methods can be transferred from one field to the other. To be more specific, can methods concerning 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 with 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 [18], [19].

Vision is used in tasks such as rope manipulation [20], [21] or cloth unfolding [22], [23], 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], [24]. Most of above-mentioned works are based on 2D vision, 3D perception of deformable objects is more challenging. Existing works employ FEM [25] or a combination of Growing Neural Gas and Particle Graph Networks [26] for better tracking the deformation. In a more recent study [27], it has been shown that deep convolutional neural network used in vision 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 the stiffness. In [28], the GelSight [29], 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 [30] present use vision to estimate the local deformation of objects during grasping, and classify objects accordingly. In [31], 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 indicating global properties of clothing according to a local sampling point. In [32] 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

Here, the main challenges 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, fusion of different sensing modalities (vision, force and tactile) may be a promising direction to pursue in future research. Also, we can consider design novel sensors for DOM.

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 the usefulness of such representation.

The object representation is used as: 1) a feedback signal for control purposes, 2) a state vector embedding the object deformation. The two usages have profound implications on choosing the type of representation.

When the representation is used as a feedback signal, one challenge is that it must 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. With vision data from different perspectives, we might be able to reconstruct the object's configuration accurately even under occlusion and noise.

When used as state variables, the challenge of getting a good representation 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. Thus, to automate this process, one promising direction maybe to learn such representation with inverse reinforcement learning.

IV. Modeling

A. Current capability

For robots to perform deformation tasks using data from various sensing modalities, we need a deformation 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. Yet, since the model should be used for control, in this section, we prefer to distinguish between local and global models – a taxonomy which 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 model is linear and can be computed in real-time with a small amount of data. Yet, since it is a local model, it should be continuously updated during task execution. Model updating methods include: Broyden rule [16], receding horizon adaption [33], local gradient descent [34], QP-based optimization [35], and Multi-armed Bandit-based methods [36]. Another advantage of the Jacobian model is that one can design a simple controller by inverting it. However, since this controller is local, it should operate via a series of intermediate target shapes [33], [35].

On the other hand, global models can be approximated with Finite Element Methods [37] 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 [38]. Moreover, they can incorporate physics models and reason about object interaction [39]. 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).

Yet, whether we use analytical or learned models, 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 [40] and [41] 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 [42] introduces 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 good simulators. While most existing robotic simulators are capable of producing rigid body kinematics and dynamic behaviours, only a fraction of them can handle deformation. A new simulator, Softgym [43] was recently proposed for benchmarking DOM, but it is currently lacking the capability of embedding robots within its environment. In the soft robotics community, [44] has recently introduced a simulator for soft robots. 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 region of validity with number of data required for training. One possible direction is to combine a simple model with a complex nonlinear model. For instance, having a linear model at lower level, and a (D)NNs learning the full model at higher level. The lower level model can be learned in few iterations to enable instant interaction between 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 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 physic-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 which combines global planning without physics simulation, with local control [45]. For an elastic object, considering its energy is another way to do planning; in this direction, Ramirez-Alpizar et al. [46] 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 [47]. More recently, [48] learns a latent representation for semantic soft object manipulation that enables (quasi) shape planning with deformable objects.

With LfD, the robot can be trained to manipulate deformable objects by an expert (usually a human). Instead of explicitly deriving a model, LfD encodes the robot trajectory and interaction force from human demonstrations [49], 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 [50].

B. Challenges and outlook

A rigid object configuration can be described in space with 6 DoF, whereas a deformable object configuration has much higher 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. It is easier to handle a deformable object by reducing its number of DoF. Examples include the use of contact points in cable harness 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, in folding, grasping at convex vertices of the clothes guarantees stability and facilitates the task [51]. Regrasp planning is highly relevant when considering tasks which 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 direction for 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 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 controllers is decided usually by the task. For instance, the authors employed a data-driven model predictive control [52] for cutting considering its predictive nature and the lower demand for manual tuning. For safe interaction in minimally invasive surgery, the authors of [53] 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 [54].

A number of 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 [55], [56], 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 [57], [58].

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, one way is to combine the offline and the online learning methods, where both are model-free. 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 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

less 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 probably in a sequential manner, to guarantee controllability.

In addition, due to the deformation during control, the contact between robot end-effector and deformable object may not always be maintained. Most existing systems commonly 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 previous sections, we centered our discussions from a scientific point of view, here, we instead discuss challenges in various applications where DOM can be translated to solutions.

Automatic laundry: A typical domestic application of DOM is laundry folding. A Tokyo-based company unveiled its prototype laundry-folding robot in 2015 (Fig. 6a). 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 [59]. Although cloth folding has been tackled in a few previous research [60]–[63], it remains largely a laboratory product (limited to structured environments, certain types of the clothes, etc). Commercializing the technology seems requiring a substantial efforts.

Assistive dressing: Due to its many challenges robot assisted dressing is still an academic endeavor. It has the potential to become an important technology not only because of the pressing needs for ageing society support. 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 variations of geometric and dynamic properties of textiles, low-cost high-reliable robot hardware, human acceptance of such technologies.

Surgical robotics: Soft tissue manipulation is mainly performed with tele-operation solely using visual feedback. Autonomous manipulation, however, still has a long way to go and demands developing various DOM hardware and software (Fig. 6d). The biggest concern for an autonomous solution is the safety of operation. A soft robot with intrinsic compliance will probably enhance the safety, but as stated in Sect. II-B and IV-B, how to model and control the interaction between soft robot and deformable is still an open problem in research.

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. 6c)) 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.

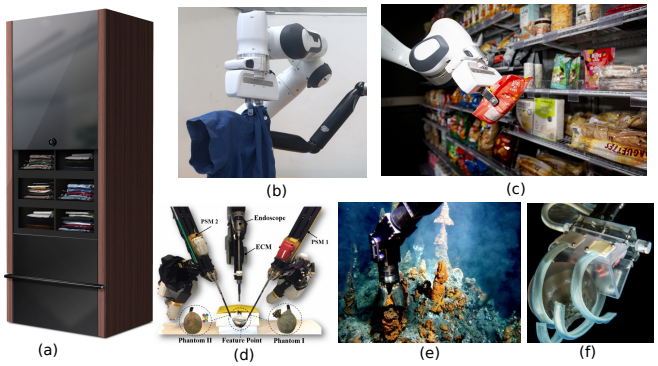


Fig. 6. Various applications of DOM – (a): laundry-folding robot from Seven Dreamers Laboratories Inc. [65], (b): A mock-up for robotics dressing assistance, (c): a robot picking a flexible bag of goods on the shelf, courtesy AIR-Lab Delft [66], (d): autonomous surgical manipulation by the dVRK system [67], (e): ROV Victor 6000 sampling black smokers (IFREMER/GENAVIR) courtesy D. Desbruyères, (f): Ultra soft underwater gripper for jellyfish [68]

Marine robotics: Underwater grasping has been led by oil and gas industry for decades, resulting in heavy machines with strong grippers for inspection and maintenance tasks (Fig. 6e). Gradually the demands turned to more detailed tasks in marine biology, sedimentology and archaeology (Fig. 6f). 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 [64].

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. A summary of challenges and ideas discussed are presented in Fig. 7.

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 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.

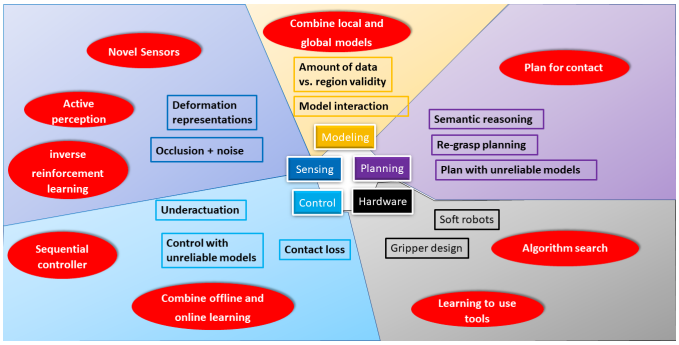


Fig. 7. A summary of open research problems and ideas/methods to pursuit discussed in this paper. Research problems in each subarea are written with bold black texts, whereas ideas/methods to resolve them are marked with bold white texts in red eclipses.

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 a 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 to react accordingly.

REFERENCES

[1] "La fontaine memory care." [Online]. Available: https://www.lafontaine-mc.com/

[2] "Cable harness." [Online]. Available: https://www.nai-group.com/products/custom-cable-harness/

[3] B. G. 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] J. 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] 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.

[16] 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.

[17] "This fruit-picking robot can pick up to 25,000 raspberries a day, and it could someday replace human workers." [Online]. Available: https://www.businessinsider.nl/robot-picks-25000-raspberries-a-day-outpaces-human-workers-2019-5/?jwsource=cl

[18] 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.

[19] 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.

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

[27] 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.

[28] 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.

[29] 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.

[30] 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.

[31] 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.

[32] 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.

[33] 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.

[34] 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.

[35] 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.

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

[37] 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.

[38] 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.

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

[40] 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.

[41] 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.

[42] 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.

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

[44] 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.

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

[46] 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.

[47] 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.

[48] 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.

[49] 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.

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

[51] 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.

[52] 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.

[53] 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.

[54] 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.

[55] 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.

[56] 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.

[57] 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.

[58] 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.

[59] 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/

[60] 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.

[61] 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.

[62] 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.

[63] 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.

[64] 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

[65] A. Wolf, "Japanese robotics biz helps lick laundry," Dec 2017. [Online]. Available: https://www.twice.com/product/japans-seven-dreamers-robot-solution-folding-laundry

[66] "The AI for retail (AIR) Lab Delft." [Online]. Available: https://icai.ai/airlab-delft/

[67] 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.

[68] 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, p. eaax5425, 2019.