2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 25-29, 2020, Las Vegas, NV, USA (Virtual)

Building 3D Deformable Object Models in Partially Observable Robotic Environments

Narmadha Sambandam[1], Stephen Rowlands[1], Angel J. Valencia[1], Pierre Payeur[1] and Ana-Maria Cretu[2]

Abstract— **Building 3D models is an important step to perform more autonomous and dexterous manipulation on deformable objects. Current techniques for modeling deformable objects are inflexible and suffer from discrepancies with actual behaviors when assumptions on their material are not properly fulfilled, or when objects can only be observed from limited viewpoints. Deformable object models require close integration between computer vision, deep learning and robotics. In this work, a framework for modeling 3D deformable objects from multi-view images is presented to support robotic manipulation.**

I. INTRODUCTION

Deformable object manipulation has become ubiquitous in numerous practical applications in the realm of robotics, gaming, and augmented reality. However, most of the dexterous manipulation tasks involving deformable objects are developed in simulated environments. This presents several integration problems when transposed to the real world, because important modeling quantities are substantially scarcer and not directly accessible as in simulators.

In this context, sensing is responsible for capturing information about changes in the object shape along with providing an appropriate representation for the extraction of modeling quantities. It also plays a significant role in handling the omnipresent challenges posed by factors such as occlusion and clutter in the scene, along with the added complications inherent to the intricate nature of deformable objects. Unlike the handling of rigid objects, robots are yet to achieve the same level of dexterity when dealing with deformable objects. This is primarily because robots must not only control the object's pose but also account for the object's changing shape [1].

II. DEFORMATION MODELING

Modeling provides robust mechanisms to formally describe the deformation behavior of the object shape. Analytical or data-driven approaches are commonly used to model deformable objects using different shape representations. These include physics models on meshes such as finite-element methods [2], mass-spring systems [3], or particles such as position-based dynamics [4]. Others use learning models on feature points such as neural networks [5].

[1] School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, Canada. {nsamb018, srowl017, avale073, ppayeur}@uottawa.ca
[2] Department of Computer Science and Engineering Université du Québec en Outaouais, Gatineau, QC, Canada. ana-maria.cretu@uqo.ca

Several challenges are faced when using deformation models in robotic environments. In principle, the configuration of these models is done through continuous feedback on the object shape. Therefore, shape estimation is an important preliminary phase that is often overlooked. The acquisition of continuous, complete and accurate information about the shape of the object being manipulated is necessary for a correct modeling and to avoid discrepancies between the representation and the actual deformation behavior, aside from the model's own limitations. Also, the model must produce reliable shape predictions when evaluated with contact information from the robotic hand. The latter is essential to generalize tasks (e.g., visual servoing [6] can more easily adapt to new tasks by modeling the deformation dynamics).

III. PARTIAL OBSERVABILITY

In most manipulation tasks, objects will only be partially observable due to sensors location and occlusions by the manipulator itself. Therefore, only partial observations are available to develop complete 3D object models. The proposed framework for 3D modeling from such partial observations is presented in Fig. 1 and detailed below.

A. MULTI-VIEW SENSING

To create a representation of deformable objects, an effective methodology for accurate and comprehensive modeling involves actively planning how the object is to be perceived from multiple viewpoints, rather than single-view perception, while capturing its deformation. However, multi-view sensing gives rise to the view selection problem, wherein the most beneficial views to achieve the highest possible accuracy are to be identified while requiring the smallest number of views. Often times, the captured data can be inadequate for building comprehensive models or too redundant, causing a degradation in performance efficiency. Next Best View (NBV) planning addresses this problem by determining the pose and settings of a vision sensor to undertake a vision task usually requiring several views [7]. The incremental view planning aspect of NBV enhances the performance of object reconstruction as RGB-D sensors capture data from multiple views on the object.

However, for manipulating deformable objects, NBV must not only select viewpoints to ensure coverage of the object but also consider the presence of the robotic gripper and its interaction with the object that impact the data collection process. Once multi-view RGB-D data is acquired from the

Workshop on Managing deformation: A step towards higher robot autonomy
International Conference on Intelligent Robots and Systems 2020, Las Vegas, NV, USA

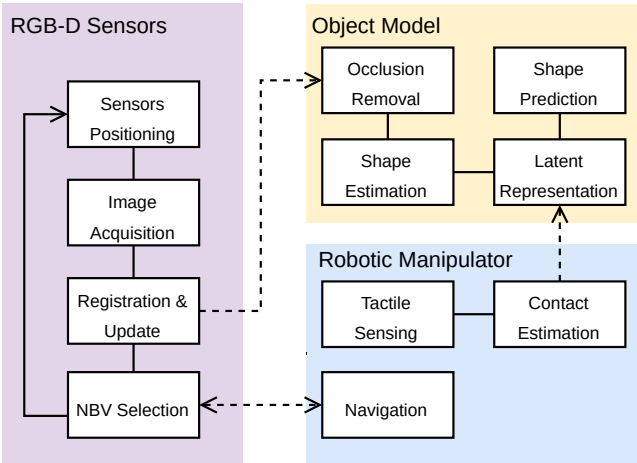

Fig. 1. A framework for modeling 3D deformable objects in partially observable robotic environments.

sensor, it must undergo registration to consolidate the data. Then, NBV planning occurs with feedback from the manipulator's pose considered in the view selection process, and an updated pose is generated toward the optimal viewpoints for the sensors to observe the object next, which is used to plan the robot navigation.

B. OCCLUSION REMOVAL

No matter how rich the sensor environment is, when manipulating 3D objects there will always be occlusions caused by the manipulator itself. Estimating the 3D shape of deformable objects during manipulation presents a unique challenge not involved when handling rigid objects as the shape will continuously change as the object is being manipulated. There exist several methods for shape completion of rigid objects including symmetry based methods [8], object detection methods, and more recently deep learning methods based on volumetric convolutional neural networks [9]. The latter rely heavily on training data to estimate the occluded parts of 3D objects. This poses additional problems for deformable objects because as they deform, algorithms are constantly presented with new data not necessarily presented in training views. This part of the work investigates dynamic geometric primitive patch-based region filling to complement acquired data where shape and deformation cannot be measured due to occlusions by the manipulator.

C. LATENT REPRESENTATIONS

But even with multi-view sensing, robots may still need to cope with missing data (e.g., objects facing corners in unknown environments) or certain physical characteristics can be very difficult to capture directly from visual measurements (e.g., material properties for objects that are non-homogeneous in composition). In these cases, operation on latent spaces turns out to be effective when handling partial observations in the data.

Latent representations are investigated here using encoder functions. In our case, the latter reduce the dimensionality of the sensing space (e.g., containing raw physical and geometric characteristics), which is transformed to a new compact space. For example, object shapes on feature points are projected to a latent space using principal component methods [10] to perform shape completion on new partial instances. In addition, to handle partial shapes, the combination of latent representations with learning models provides a potential advantage to generalize shape predictions with classes of objects outside the training data since the dynamic behavior between shape and gripper contacts are transferable in latent representations.

IV. CONCLUSIONS

This abstract overviews original research work pursued to tackle 3D modeling of deformable objects undergoing robotic manipulation while being observed by multiple RGB-D sensors. The presentation will put forth early experimental results and provide an opportunity to exchange on this challenging topic.

REFERENCES

[1] 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, Nov. 2018, doi: 10.3390/robotics7040074.

[2] A. Petit, V. Lippiello, and B. Siciliano, "Real-time tracking of 3D elastic objects with an RGB-D sensor," In IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (Hamburg, Germany: IEEE), pp. 3914–3921, 2015, doi:10.1109/IROS.2015.7353928

[3] L. Zaidi, J. A. Corrales, B. C. Bouzgarrou, Y. Mezouar, and L. Sabourin, "Model-based strategy for grasping 3D deformable objects using a multi-fingered robotic hand," Robotics and Autonomous Systems, 95, pp. 196–206, 2017, doi:10.1016/j.robot.2017.06.011

[4] P. Güler, K. Pauwels, A. Peropan, H. Kjellstroöm, and D. Kragic, "Estimating the deformability of elastic materials using optical flow and position-based dynamics," In IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids) (Seoul, South Korea: IEEE), pp. 965–971, 2015, doi:10.1109/HUMANOIDS.2015.7363486

[5] A. Cretu, P. Payeur and E. M. Petriu, "Soft Object Deformation Monitoring and Learning for Model-Based Robotic Hand Manipulation," in IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 42, no. 3, pp. 740-753, June 2012, doi: 10.1109/TSMCB.2011.2176115.

[6] A. Cherubini, J. Leitner, V. Ortenzi, and P. Corke, "Towards vision-based manipulation of plastic materials," In IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (Madrid, Spain: IEEE), pp. 485–490, 2018, doi:10.1109/IROS.2018.8594108

[7] C. McGreavy, L. Kunze, and N. Hawes, "Next best view planning for object recognition in mobile robotics," CEUR Workshop Proceedings, pp. 1–9, 2017.

[8] D. Shiebener, A. Schmidt, N. Vahrenkamp and T. Asfour, "Heuristic 3D object shape completion based on symmetry and scene context" In IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (Daejeon, South Korea: IEEE), pp. 74-81, 2016, doi:10.1109/IROS.2016.7759037

[9] J. Varley, C. DeChant, A. Richardson, J. Ruales and P. Allen, "Shape Completion Enabled Robotic Grasping," In IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (Vancouver, Canada), pp. 2442-2447, 2017, doi:10.1109/IROS.2017.8206060

[10] D. Rodriguez, C. Cogswell, S. Koo and S. Behnke, "Transferring Grasping Skills to Novel Instances by Latent Space Non-Rigid Registration," In IEEE International Conference on Robotics and Automation (ICRA), Brisbane, QLD, pp. 4229-4236, 2018, doi: 10.1109/ICRA.2018.8461169.