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

SOMA: A Data-Driven Representation Framework for Semantic Soft Object Manipulation

Peng Zhou[a], Jihong Zhu[b] and David Navarro-Alarcon[a]

***Abstract*— Soft object manipulation has recently received a lot of attention from the robotics community due to its vast potential applications. Most existing vision-based methods are case-specific, as their representation algorithms typically rely on "hard-coded" features to characterize the object's shape. In this paper, we present SOMA, a new feedback representation framework for Semantic Soft Object MAnipulation. We introduce internal automatic representation layers between a low-level geometric feature extraction and a high-level semantic shape analysis. Thereby, allowing to identify the semantic function of each compressed feature and form a valid shape classifier. The high-level semantic layer shows how to perform (quasi) motion planning shaping tasks for soft objects. In this way, this decomposed framework makes soft object representation more generic and scalable. To validate the proposed methodology, we report a detailed experimental study with bimanual manipulation tasks.**

I. Introduction

Studies have shown that the manipulation of soft objects is a crucial and indispensable to achieve high autonomy in robot manipulators [1]. Although great progress has been recently achieved, the feedback manipulation of soft objects is still largely an open research problem. The implementation of these types of advanced manipulation capabilities is complicated by various issues, amongst the most important is the difficulty in characterizing the feedback shape of a soft object. Our aim in this work is to develop new data-driven methods that can quantitatively represent deformable shapes.

Several "classical" methods demonstrated how feedback controls can be used to deform a soft object into a desired 2D shape based on geometric features e.g. angles, curvatures, catenaries [2]; Their disadvantage is that they are case-specific, thus, can only be used to perform a single shaping action. Some works have addressed this issue by developing generic representations that only require sensory data [3]. These methods, however, create very large feature vectors, which might not be the most efficient feedback metric. A useful approach is to automatically compute generic feedback features (e.g. as in direct visual servoing [4]) and combine them with dimension reduction techniques, as in e.g. [5]. Latest applications have also examined attribute-based approaches [6]. Our approach has the same purpose, with an emphasis on the combination of shape analysis and semantic attributes to achieve a comprehensive semantic shape analysis of shape deformation.

In this paper, we present a general data-driven representation framework for semantic soft object manipulation. As shown in Fig. 1, we create a three-level representation, which involves a low-level soft object geometrical shape processing, a mid-level data-driven representation learning, and a high-level semantic soft shape analysis. With this semantic shape knowledge, we can foresee the possibility of shape deformation and provide useful suggestions for guiding the soft manipulation. One of the most important scenarios is to achieve shape planing for soft objects in the reduced feature space. Such decomposed and semantic framework will provide additional flexibility and intelligence in the design of complicated soft manipulation systems. The main contributions of this work are summarized as follows:

- An effective representation framework for soft object analysis in manipulation tasks.
- A set of novel semantic analysis approaches for high-level shape deformation analysis.

II. Framework Overview

SOMA exhibits a possibility of constructing valid and compact features to represent soft objects without $priori$ information. Based on the built data-driven pipeline, semantic analysis techniques can be introduced into a robotic manipulator to datamine explainable deformation knowledge for guiding soft manipulation tasks.

Low-Level Processing: As the first step in SOMA, this level provides the low-level extracted geometric features. These low-level features include the contour descriptor, surface normal, and boundary segments. However, lacks of representation efficiency and hierarchical structure preclude us from using raw low-level features of soft objects to its fullest extent. Consequently, We should reduce the dimensionality of the low-level extracted shape features.

Mid-Level Representation: As shown in Fig. 1, two modularized data-driven learning phases are applied to produce a set of compressed and semantic representations. In Phase I, we employ both linear and non-linear dimensionality reduction modules to compress low-level features. (Linear model: principal component analysis (PCA), the non-linear model: the auto-encoder network). The entire phase I outputs prime and concise representations of soft objects. In Phase II, we design a semantic feature analysis algorithm to detect the function of each reduced feature (the corresponding result is shown in Fig. 2). In addition, we employ k-nearest neighbor (kNN) to classify each shape into one of the predefined soft

[a]The authors are with The Hong Kong Polytechnic University, Kowloon, Hong Kong.

[b]Jihong Zhu with Delft University of Technology, Delft, Netherlands.

This work is supported in part by the Research Grants Council (RGC) of Hong Kong under Grant Number 14203917. Corresponding author e-mail: jeffery.zhou@connect.polyu.hk

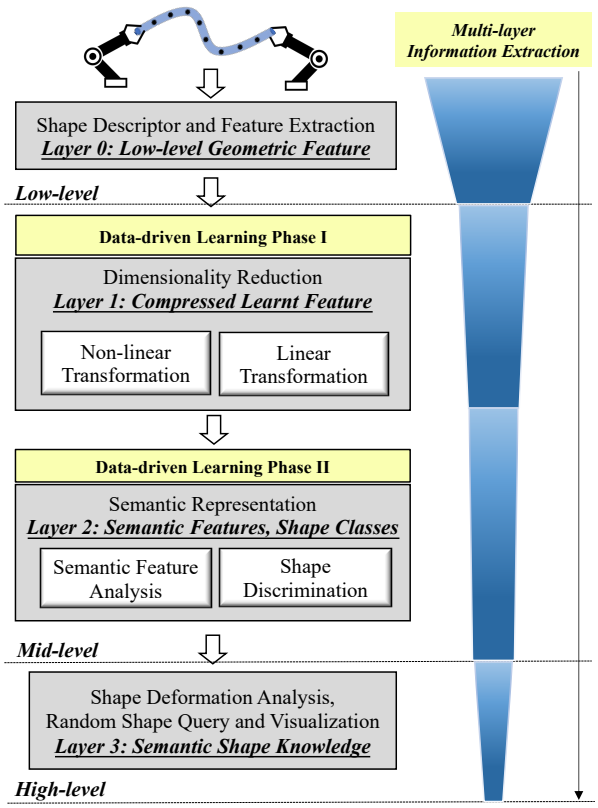

Fig. 1. Conceptual representation of the proposed framework — SOMA that fully describes and represents the soft objects for bimanual manipulation tasks from different layers.

object shape categories, such as arch class, helix class and s-shaped class.

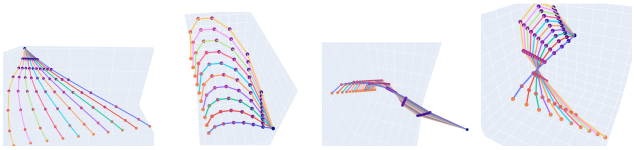

Fig. 2. The reconstructed shapes from the different semantic features.

High-Level Analysis: Soft object deformation knowledge plays an important role in manipulation tasks. This knowledge comprises two parts, namely, shape deformation relations and shape morphing process. We use semantic compressed features to construct a 3D visualization configuration for observing the deformation relations between any pair of different shape categories. Subsequently, as shown in Fig. 3, based on the known semantic compressed features, shape morphing process can show an overview of the shape deformation trajectory in the compressed feature space divided by different shape categories.

III. Conclusion

In this paper, we present a generic 4-layer data-driven representation framework for soft objects in bimanual manipulation tasks. Such decomposed learning phases can improve flexibility and robustness for the soft object analysis in bimanual manipulation tasks. The experimental results show the viability of using this framework for different levels of representation. Future studies need to test point cloud data for different soft objects.

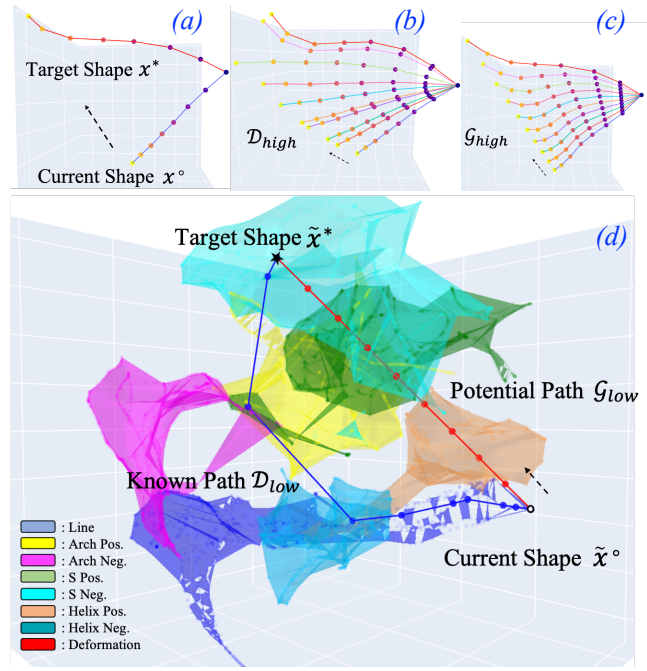

Fig. 3. Visualization of the algorithm operations designed for a shape planing scenario, with (a) showing the beginning shape x° and the target shape x^*, and (b) and (c) present the real shape morphing processes $\mathcal{D}_{high}$ and $\mathcal{G}_{high}$ based on the shortest path searching on the known shape set and the potential shape set, respectively. (d) presents their corresponding deformation traces in a 3D shape space.

References

[1] H. B. Amor, A. Saxena, N. Hudson, and J. Peters, "Special issue on autonomous grasping and manipulation," *Auton. Robots*, vol. 36, no. 1-2, pp. 1–3, 2014.

[2] D. Navarro-Alarcon, H. M. Yip, Z. Wang, Y.-H. Liu, F. Zhong, T. Zhang, and P. Li, "Automatic 3-d manipulation of soft objects by robotic arms with an adaptive deformation model," *IEEE Trans. Robot.*, vol. 32, no. 2, pp. 429–441, 2016.

[3] J. Zhu, B. Navarro, P. Fraisse, A. Crosnier, and A. Cherubini, "Dual-arm robotic manipulation of flexible cables," in *IEEE/RSJ Int. Conf. on Robots and Intelligent Systems*, 2018, pp. 479–484.

[4] E. Marchand, "Subspace-based direct visual servoing," *IEEE Robot. Autom. Lett.*, vol. 4, no. 3, pp. 2699–2706, 2019.

[5] K. M. Digumarti, B. Trimmer, A. T. Conn, and J. Rossiter, "Quantifying dynamic shapes in soft morphologies," *Soft Robot.*, 2019.

[6] P.-Y. Laffont, Z. Ren, X. Tao, C. Qian, and J. Hays, "Transient attributes for high-level understanding and editing of outdoor scenes," *ACM Trans. Graph.*, vol. 33, no. 4, pp. 1–11, 2014.