Accurate Mesh Initialization for Deformable Object Tracking with Gaussian Splatting

Alberta Longhini, Marco Moletta, Michael Welle
Primary contact: albertal@kth.se
Division of Robotics, Perception and Learning, Sweden, 2024

Introduction

State estimation of deformable objects remains a persistent challenge in robotics. Unlike rigid objects, defining a canonical state representation, such as the object's pose, remains difficult and is often dependent on the specific task to be accomplished by the robot. Particle-based representations, such as graphs or landmark points, are commonly used due to their interpretability and strong alignment with analytical models of deformable objects like cloth (2). However, the frequent self-occlusions in crumpled configurations of deformable objects pose significant challenges for accurate state estimation of particle representations, making it difficult to consistently capture and track the object's state.

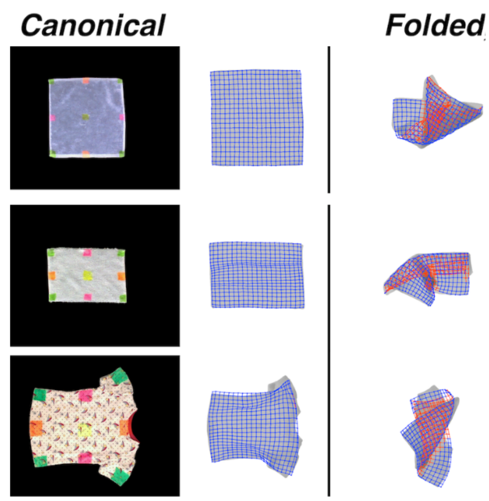

Figure 1: State estimation of cloths in crumpled configurations. Image taken from (1)

Current state estimation methods for cloth-like deformable objects, based on raw observations such as point clouds and images, often rely on graph-based representations due to their ability to encode structural information and their computational efficiency when paired with Graph Neural Networks (3). Gaussian Splatting (4) has emerged as a promising approach for tracking graph representations of cloths. However, current solutions require an initial mesh configuration as input, typically from a flattened state, to begin tracking accurately (5). This reliance on precise mesh initialization poses a limitation when the initial configuration is crumpled. To address this, state estimation methods from static configurations play a crucial role by providing an initial estimate in crumpled configurations. These methods involve inferring the cloth's state either by learning a connectivity network over observed point clouds (6) or by reconstructing the state using template-matching techniques (1). While combining the static estimation could enhance both accuracy and robustness in tracking deformable objects with Gaussian Splatting techniques, this integration remains largely unexplored.

This project aims to investigate the integration of static state estimation methods to improve the accuracy and robustness of state estimation and tracking for deformable objects. Specifically, we propose leveraging static estimation to initialize or refine mesh representations used in Gaussian Splatting during dynamic tracking. We hypothesize that by combining these two approaches, particle-based state estimation methods can achieve greater accuracy, particularly in long-horizon tasks where self-occlusions of cloth accumulate, posing significant challenges.

Additionally, we speculate that this combination could pave the way for active learning techniques, where the robot performs specific manipulation actions to maximize the refinement of its state estimation. More specifically, when the initial static mesh is imprecise, tracking it over time could reveal inconsistencies, such as nodes moving in unexpected ways compared to their neighbors. These discrepancies could signal errors in the current graph representation, enabling the system to select actions that efficiently refine the mesh and improve the accuracy of the cloth's state representation.

Goal and Objectives

The goal of this project is to develop a robust method for state estimation and tracking of deformable objects by integrating static state estimation techniques with Gaussian Splatting. This approach aims to enhance the accuracy of particle-based representations, particularly in long-horizon tasks involving crumpled configurations and self-occlusions. If the results are positive, we expect to publish this work in a top-tier robotics or machine learning conference. Your tasks will be to:

- Conducting a thorough literature review to identify gaps in current research on deformable object state estimation and tracking.
- Familiarizing with graph-tracking frameworks already available in our lab, such as Gaussian Splatting, both in simulation and the real world.
- Implementing baseline models and collecting demonstration datasets in both simulated and real environments.
- Designing a method for evaluating real-world tracking performance where ground truth labels are not available.
- Testing and evaluating the proposed hypothesis in both simulated and real-world experiments comparing the proposed method against the baselines.
- Report the results in a scientific manner.

Requirements

This project is suitable for students interested in the field of robotics and machine learning. You should have strong programming skills in any high-level language (Python, C++, Julia, etc..). Some basic knowledge about ROS is fundamental, and some experience with training machine learning models is needed. Desirable skills include knowledge knowledge of / experience with robotic manipulation, optimizing code for speed and parallelism, and experience with robotic hardware.

References

[1] Wenbo Wang, Gen Li, Miguel Zamora, and Stelian Coros. Trtm: Template-based reconstruction and target-oriented manipulation of crumpled cloths. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pages 12522–12528. IEEE, 2024.

[2] Alberta Longhini, Yufei Wang, Irene Garcia-Camacho, David Blanco-Mulero, Marco Moletta, Michael Welle, Guillem Alenyà, Hang Yin, Zackory Erickson, David Held, et al. Unfolding the literature: A review of robotic cloth manipulation. *arXiv preprint arXiv:2407.01361*, 2024.

[3] Alvaro Sanchez-Gonzalez, Jonathan Godwin, Tobias Pfaff, Rex Ying, Jure Leskovec, and Peter Battaglia. Learning to simulate complex physics with graph networks. In *International conference on machine learning*, pages 8459–8468. PMLR, 2020.

[4] Guanjun Wu, Taoran Yi, Jiemin Fang, Lingxi Xie, Xiaopeng Zhang, Wei Wei, Wenyu Liu, Qi Tian, and Xinggang Wang. 4d gaussian splatting for real-time dynamic scene rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 20310–20320, 2024.

[5] Alberta Longhini, Marcel Büsching, Bardienus Pieter Duisterhof, Jens Lundell, Jeffrey Ichnowski, Mårten Björkman, and Danica Kragic. Cloth-splatting: 3d cloth state estimation from rgb supervision. In *8th Annual Conference on Robot Learning*.

[6] Xingyu Lin, Yufei Wang, Zixuan Huang, and David Held. Learning visible connectivity dynamics for cloth smoothing. In *Conference on Robot Learning*, pages 256–266. PMLR, 2022.