A compact cloth state representation that enables generalization across different cloth shapes

Jay Kamat, Júlia Borràs, Carme Torras

***Abstract*— Cloth manipulation is a difficult problem mainly because of the high dimensionality of its configuration space, which makes a simplified representation of deformation essential to enable reasoning on cloth state. In this extended abstract, we discuss the novel ideas proposed in our previous work [1]. First, we discuss the *dGLI disk representation*, a circular heatmap constructed from the topological indices, dGLI, computed pairwise for edge segments on the cloth border, arranged on a circular grid. This heatmap uncovers patterns that correspond to features of the cloth state that are consistent for different shapes, sizes and positions of the cloth. We then introduce *Cloth StatE representation* (CloSE), a representation obtained by abstracting the corners and the first fold locations from the dGLI disk onto a circle. This representation is compact, while still preserving the invariances observed on the dGLI disk. Finally, we show the strengths of the CloSE representation in two relevant applications: semantic labeling and high- and low-level planning for single folds, as well as preliminary results suggesting a smooth extension to several folds. The reader can find extended information on the paper website: `https://jaykamat99.github.io/close-representation/`**

I. Introduction

As opposed to rigid objects or articulated objects, the configuration space of cloth is high dimensional, making state estimation very complex. Several reviews in the literature point to the need of a simplified representation that could pave the way for more efficient learning methods [2–4].

In recent years, many datasets with both real and simulated clothes have appeared, as summarized in [5]. Real images lack ground truth information on the configuration of the cloth, and many reconstruction works [6, 7] rely on realistic simulated images where the ground truth is the full mesh. But even then, there is no real understanding of the configuration of cloth other than comparing to predefined states or recognizing which pixels are sleeves or collars. Existing solutions for navigating between cloth states use zenithal RGB-D images of folded configurations [8, 9], learning transitions between observed silhouettes to construct a graph in a learned latent space. However, this space is not continuous and the learned graph is specific to the training examples. In this abstract, we discuss a cloth representation that is general to any cloth shape and is also continuous.

This abstract summarizes our work in [1] where we apply the derivative of a topological measure, the Gauss Linking Integral (GLI) [10–12], to the cloth border. We define cloth border as the silhouette curve corresponding to the flat, unfolded state of a cloth. Following from our previous works [13, 14], we compute the dGLI matrix, which contains the derivative of GLI (dGLI) between any two segments on the cloth border. The dGLI between two segments is a close-form formula easy-to-compute. In our work we propose to rearrange this matrix in a circular grid, calling it the dGLI disk. This dGLI disk uncovers interesting geometric structures when the cloth is folded that are invariant for

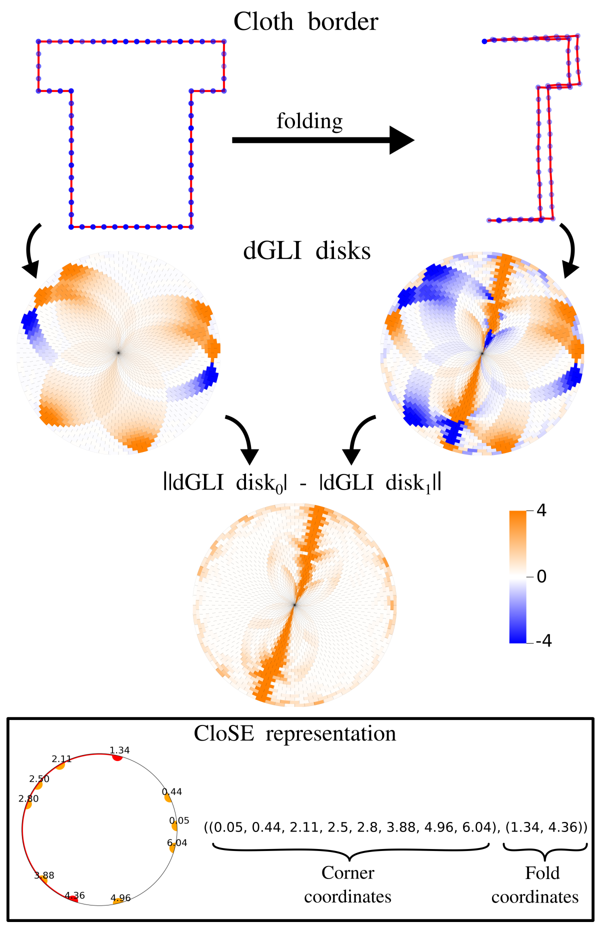


Fig. 1: Example of CloSE derivation: From the border, we compute the dGLI disk where each petal corresponds to a corner. The difference between dGLI disks for the start and end frames allows to extract the fold information, so as to obtain the final CloSE descriptor that is shown graphically and numerically at the bottom of the figure.

The research has been partially supported by project PID2023-152259OB-I00 (CHLOE-MAP), MCIU/ AEI /10.13039/501100011033, project ROBassist (CSIC code 202450E060), and by ERDF, UE and project SoftEnable (HORIZON-CL4-2021-DIGITAL-EMERGING-01-101070600)

All authors are with the Institut de Robòtica i Informàtica Industrial, CSIC-UPC

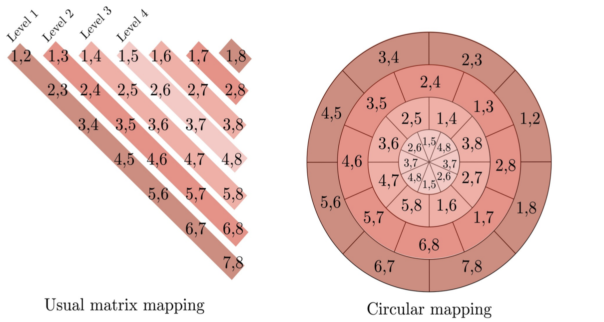


Fig. 2: The Circular grid mapping in comparison to the semi-matrix mapping

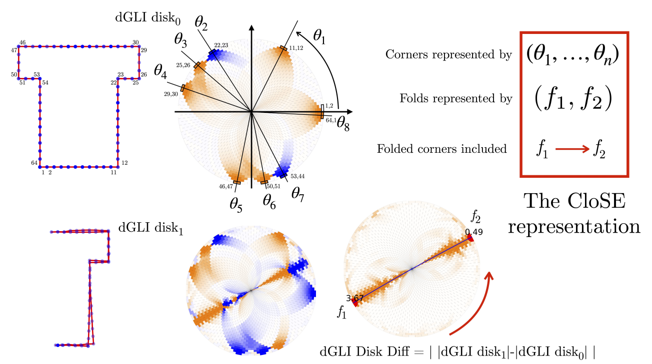


Fig. 4: Computation of the CloSE representation from the dGLI disks

different shapes, sizes, locations and orientations of the cloth. From the dGLI disk, we derive the CloSE representation. A representation that is compact, continuous and allows to reconstruct cloth state from it. It encodes:

- The shape of the cloth, i.e., the location of the corners of the cloth border on the dGLI disk.
- The location of the fold and its orientation (which side folds up)

Fig. 1 shows the starting and final cloth borders of a folding action, their corresponding dGLI disk heatmaps, the subtraction operation that allows to locate the fold, and the CloSE representation abstracted from the above dGLI disks.

II. The dGLI Disk

The dGLI disk is a circular reformulation of the dGLI matrix proposed in [13] and in other papers on the Gauss Linking Integral [10–12] that usually utilize the matrix form. The mapping from the semi-matrix to the disk is illustrated in Fig.2. This new representation offers a powerful way to visualize key information such as corner locations, the presence of folds, and their positions (Fig.3c).

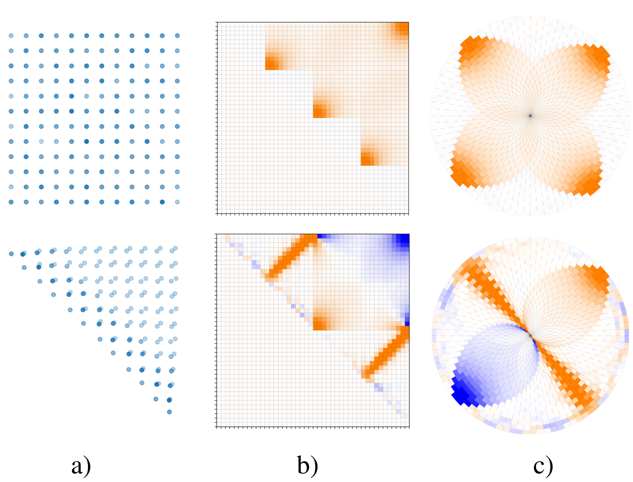


Fig. 3: a) A square napkin mesh. b) The corresponding original dGLI matrix as in [13] and c) Our proposed arrangement in the dGLI disk

Importantly, because this representation solely depends on the relative positions of border segments, it is invariant to the cloth's orientation and position in space. We depict the dGLI disks in our figures as heatmaps where shades of orange correspond to positive and blue to negative values. The dGLI disk in Fig.3c [top] represents the cloth in the unfolded state and the dark petal-like structures indicate the corners, while the dGLI Disk in Fig.3c [bottom] represents the folded cloth. Here, the curve appearing alongside the petal-like shapes reveals the fold line. The petals with inverted colors (signs/dGLI values) indicate which corners have been folded. You can find additional examples on the paper website.

III. The Cloth StatE (CloSE) representation

From the dGLI disk, we abstract out 1) the location of the corners, 2) the fold location and 3) information about which corners have been folded. This representation, the CloSE representation, is compact and continuous, while maintaining the invariances from the dGLI disk.

Following the notation in Fig. 4, the CloSE representation for a cloth state can be represented as

$$((\theta_1, \ldots, \theta_n), (f_1, f_2)),$$

where n is the number of corners. By construction, each θ_i and f_i are numbers from $[0, 2\pi)$, because they are angles in radians. This information is sufficient to decipher the state of the cloth. Furthermore, since the distance correspondence along the border is proportional to the distances in the circle, given the initial cloth border and any CloSE representation, we can reconstruct the folded state of the cloth. Fig.5 presents the CloSE representation alongside the reconstructed folded border.

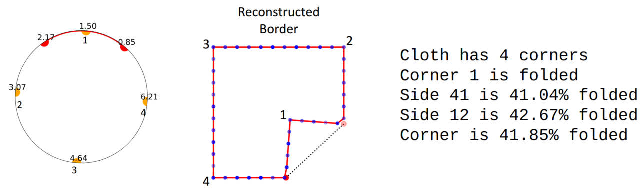
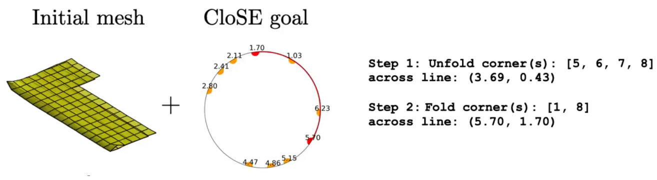
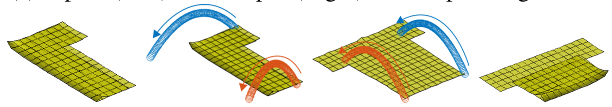


Fig. 5: CloSE, Reconstruction and Semantic Label



(a) Inputs (Left) and Output (Right) for the planning Case 2



(b) Fold Visualization - Case 2

Fig. 6: Planning outputs

The CloSE representation is continuous in the sense that moving the fold endpoints (f_1, f_2), represented as red dots in Fig.5 - left, along the circle moves the fold in a continuous way. The visualization of this property is presented in the video that the reader can find on the website.

A. Semantic Labeling

From the CloSE representation, we can identify the semantic state of cloth by interval reasoning on the circle. Each fold is an interval (f_1, f_2). The corners, θ_i, that lie inside this interval are the folded corners. In addition, each f_i falls inside an edge of the cloth, that is

$$\theta_j < f_i < \theta_{j+1} \text{ for any } i \in 1, 2, j \in 1 \dots n,$$

The angular distances from f_i to its neighbors θ_i and θ_{i+1} are proportional to the position of the fold along the folded edge at the corresponding corner. Therefore, we can also identify where the fold is on the folded edge and define semantic labels depending on their relative interval positions.

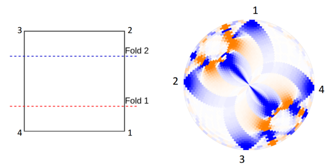
One example of the obtained labels is shown in Fig. 5. More examples and the code to generate the labels can be found on our website. Note that we obtain these without any learning, by just reasoning on our CloSE representation.

B. Planning

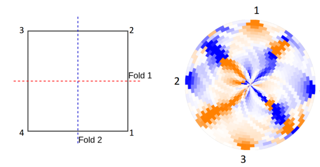
Given the initial border of the cloth, and the CloSE representation of the desired fold state, we can plan a sequence of manipulations to provide the semantic (high-level) instructions and the low-level trajectory of the corners to manipulate.

Since the CloSE representation tracks the folds, one can reason over the cloth's current state and plan the high-level intermediate states. Fig. 6 shows one of the two cases that we encounter when we are dealing with one fold. In case 1, both the initial and final configurations are in the same semantic region, i.e., the fold encloses the same corner(s). Then, we can directly manipulate the folded corner(s) to their goal position(s).

In case 2, the initial and the goal configurations are in different semantic states, so we need to take a multi-step approach to reach the goal. Note that for single fold examples, the unfolded state is the intermediate state.



(a) Independent double folds



(b) Dependent double folds

Fig. 7: Double folds

For low-level planning, that is, deciding where to pick the cloth and where to place it, the CloSE representation proves useful. In the cases where there are more than 2 folded corners that require to be moved, we can pick 2 corners such that the area of the trapezoid formed between such corners and the fold points on the border is maximum. The idea here is to control the maximum possible area of the cloth. Our manipulation actions are shown in Fig. 6b. We also run the trajectories in simulation and these can be seen in the video on the website.

IV. Discussion and Future Work

The CloSE representation derived from the dGLI disk is compact, continuous, and is general to any cloth shape. Since the dGLI disk is generated by considering only the relative positions of the border edges, the dGLI disk and hence the CloSE representation is invariant to the size, location and orientation of the cloth. We also prove the advantages of this general, compact representation on two relevant tasks: semantic labeling and high- and low- level manipulation planning for folding.

The results discussed in this extended abstract are restricted to a single fold, however, preliminary results show that similar fold patterns emerge on the dGLI disk when we perform more than one fold. Fig.7 shows two examples, with the schematic of the folds performed on the cloth along side its final dGLI disk. When there is more than one fold we encounter two cases: dependent or independent folds. Independent folds occur in separate parts of the cloth and the presence of either fold does not affect the other, while in the dependent case, the first fold is affected by the second fold. i.e., we cannot unfold the first fold before unfolding the second. We are working on an extension to the CloSE representation that captures it seamlessly. Future work needs to be done in this direction to build a better CloSE representation for other type of folds and wrinkles.

This approach is interesting for several reasons. A direct application enables automatic labeling of datasets and reasoning on the configuration space of cloth in simulation. In addition, it opens the door for learning methods to be able to generalize across different cloth types that have never been observed before. We believe the path towards general data-driven approaches requires the adoption of such powerful representations.

REFERENCES

[1] J. Kamat, J. Borràs, and C. Torras, "Close: A compact shape- and orientation-agnostic cloth state representation," *arXiv preprint arXiv:2504.05033*, 2025.

[2] H. Yin, A. Varava, and D. Kragic, "Modeling, learning, perception, and control methods for deformable object manipulation," *Science Robotics*, vol. 6, no. 54, p. eabd8803, 2021.

[3] 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," *International Journal of Robotic Research*, vol. 37, no. 7, pp. 688–716, 2018.

[4] J. Zhu, A. Cherubini, C. Dune, D. Navarro-Alarcon, F. Alambeigi, D. Berenson, F. Ficuciello, K. Harada, J. Kober, X. Li, *et al.*, "Challenges and outlook in robotic manipulation of deformable objects," *IEEE Robotics & Automation Magazine*, vol. 29, no. 3, pp. 67–77, 2022.

[5] A. Longhini, Y. Wang, I. Garcia-Camacho, D. Blanco-Mulero, M. Moletta, M. Welle, G. Alenyà, H. Yin, Z. Erickson, D. Held, *et al.*, "Unfolding the literature: A review of robotic cloth manipulation," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 8, 2024.

[6] J. Bednarik, P. Fua, and M. Salzmann, "Learning to reconstruct texture-less deformable surfaces from a single view," in *Int. Conf. on 3d vision (3DV)*, pp. 606–615, 2018.

[7] C. Chi and S. Song, "Garmentnets: Category-level pose estimation for garments via canonical space shape completion," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 3324–3333, 2021.

[8] M. Lippi, P. Poklukar, M. C. Welle, A. Varava, H. Yin, A. Marino, and D. Kragic, "Enabling visual action planning for object manipulation through latent space roadmap," *IEEE Transactions on Robotics*, vol. 39, no. 1, pp. 57–75, 2022.

[9] D. Tanaka, S. Arnold, and K. Yamazaki, "Emd net: An encode–manipulate–decode network for cloth manipulation," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 1771–1778, 2018.

[10] S. L. Ho, *Topology-based character motion synthesis*. PhD thesis, University of Edinburgh, 2011.

[11] F. T. Pokorny, J. A. Stork, and D. Kragic, "Grasping objects with holes: A topological approach," in *2013 IEEE international conference on robotics and automation*, pp. 1100–1107, IEEE, 2013.

[12] D. Zarubin, V. Ivan, M. Toussaint, T. Komura, and S. Vijayakumar, "Hierarchical motion planning in topological representations," *Proceedings of Robotics: Science and Systems VIII*, 2012.

[13] F. Coltraro, J. Fontana, J. Amorós, M. Alberich-Carramiñana, J. Borràs, and C. Torras, "A representation of cloth states based on a derivative of the gauss linking integral," *Applied Mathematics and Computation*, vol. 457, p. 128165, 2023.

[14] J. Borràs, A. Boix-Granell, S. Foix, and C. Torras, "A virtual reality framework for fast dataset creation applied to cloth manipulation with automatic semantic labelling," in *IEEE International Conference on Robotics and Automation (ICRA 2023)*, pp. 0210031–0210311, IEEE, 2023.