

A Representation of Cloth States based on a Derivative of the Gauss Linking Integral

Franco Coltraro^{a,*}, Josep Fontana^b, Jaume Amorós^b, Maria Alberich-Carramiñana^{a,b}, Júlia Borràs^a, Carme Torras^a

^a*Institut de Robòtica i Informàtica Industrial, CSIC-UPC,
C/ Llorens i Artigas 4-6, 08028, Barcelona, Spain.*

^b*Departament de Matemàtiques, Universitat Politècnica de Catalunya, Barcelona, Spain.*

Abstract

Robotic manipulation of cloth is a complex task because of the infinite-dimensional shape-state space of textiles, which makes their state estimation very difficult. In this paper we introduce the dGLI *Cloth Coordinates*, a finite low-dimensional representation of cloth states that allows us to efficiently distinguish a large variety of different folded states, opening the door to efficient learning methods for cloth manipulation planning and control. Our representation is based on a directional derivative of the *Gauss Linking Integral* and allows us to represent spatial as well as planar folded configurations in a consistent and unified way. The proposed dGLI Cloth Coordinates are shown to be more accurate in the representation of cloth states and significantly more sensitive to changes in grasping *affordances* than other classic shape distance methods. Finally, we apply our representation to real images of a cloth, showing that with it we can identify the different states using a distance-based classifier.

Keywords: semantic state labelling, robotic cloth manipulation, deformable object representation and classification, Gauss Linking Integral (GLI)

1. Introduction

Textile objects are important and omnipresent in many relevant scenarios of our daily lives, like domestic, healthcare, or industrial contexts. However,

*Corresponding author

Email address: franco.coltraro@upc.edu (Franco Coltraro)

as opposed to rigid objects, whose pose is fixed with position and orientation, textile objects are challenging to handle for robots because they change shape under contact and motion, resulting in an infinite-dimensional configuration space (when considered as continuous surfaces in 3D space). This huge dimensional jump makes existing perception and manipulation methods difficult to apply to textiles. Recent reviews on cloth manipulation, like [28, 35], agree on the need to find a simplified representation that enables more powerful learning methods to solve different problems related to cloth manipulation.

Different representations have been used in the literature of cloth manipulation, e.g. silhouette representations [22] or contours [9], assuming the high-level reasoning on cloth states was given. More modern end-to-end learning approaches use RGB-D images directly [20, 29, 32], but only very simple actions can be defined due to the limited state representation. In addition, these deep-learning based methods need large amounts of real or simulated data (e.g. [17, 21]) that are expensive to obtain and label, as no underlying previous knowledge is used to understand the geometric relationship between different states. Therefore, finding a low-dimensional representation for cloth based on low-level features remains an active open problem, while the high-level aspect of understanding cloth deformation is still almost unexplored.

Furthermore, to enable reasoning, abstraction and planning, rigid object manipulation applies object recognition methods in order to link objects to actions [6, 34]. Contacts among objects are estimated to recognize states such as “on top of”, “inside of” [2]. However, when it comes to cloth manipulation, no work has explored the semantic state identification that could lead to particular actions depending on the task in mind. For simpler deformable objects like a box with an articulated lid, the open configuration clearly allows the action of closing the box or picking something from inside. An equivalent example for cloth would be to recognize a folded corner that needs to be either flattened back if the task is to lay the cloth flat on the table, or picked up if the task is folding. In this context, we wish to decompose the configuration space of a piece of cloth into macro-states (or just states), where each state is the set of cloth configurations that can be manipulated in the same way, i.e., that have similar grasping affordances.

In this work, we present a coordinate representation of the configuration of cloth as an upper triangular matrix form (see Figure 1). This representation can be computed with a closed-form formula from low-level features

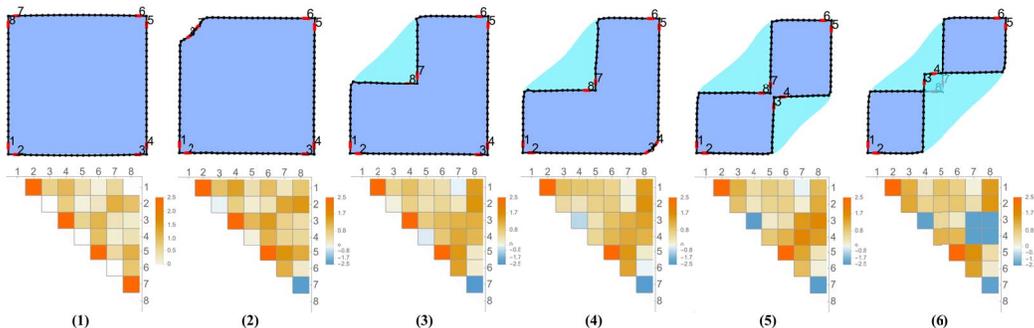


Figure 1: Folding sequence of a quadrangular cloth with its associated dGLI *cloth coordinates*, represented as upper triangular matrices. Each matrix element m_{ij} is a geometrical value corresponding to the *dGLI* between the segments i and j highlighted in red in the corresponding folded state of the cloth. Notice how some values of the matrix change sign when corners are folded or cross each other.

of the cloth –the position of its boundary– and enables the recognition and classification of high-level states, since we can define a distance between cloth configurations. This representation allows us to classify different configurations into states that we identify as “different”, meaning that they afford different actions.

The fact that our low-dimensional representation relies only on the position of the cloth’s boundary is supported by the result that under certain conditions the boundary curves of a textile determine completely its position in space. Indeed, in [3] it is proven that a generic simple, closed, piecewise regular curve in space can be the boundary of only finitely many developable surfaces (i.e. Gaussian curvature 0) with non-vanishing mean curvature. Since the original state of a piece of cloth is unfolded and flat, the set of possible states, if we assume that the textile is inextensible (i.e. constant first fundamental form in time, see [7]) is precisely the set of developable surfaces isometric to a fixed one. For problems such as the study of cloth dynamics it is not necessary for the boundary problem to have a unique solution. It suffices to know that it will always have a finite set of solutions, because this solution set is then discrete, with different solutions separated by a nontrivial jump in any tagging energy, local coordinates, etc. This implies that during a continuous cloth motion, the position of a garment is determined by the location of its boundary.

Our coordinates are based on a topological index, the *Gauss Linking Integral (GLI)*. This index has been used in the past for robotic manipulation

[16, 24, 30, 31, 37] but can only be applied to 3D curves. For a pair of almost co-planar curves, as the boundary curves of a folded garment, the *GLI* vanishes and ceases to be informative. In order to consider almost co-planar curves (as well as 3D curves), we introduce in this work the concept of the directional derivative of the *GLI*, *dGLI*, applied to a pair of curves. The *dGLI* is symmetric on the curves and it only depends on the relative position between them. We assign the *dGLI Cloth Coordinates* to a state of a garment as follows: first select a subset of edges (it may contain the whole of them) from a discretization of the boundary of the garment; then fix an ordering on these edges and compute the *dGLI* between any pair of edges in their spatial position of the current state of the garment; this gives a symmetric matrix from which only the upper triangular part is taken in order to avoid redundancies; the *dGLI cloth coordinates* of the state are precisely the entries of this upper triangular matrix (see Figure 1). Our resulting representation can be computed efficiently and is invariant under isometric movements of the garment (i.e. rotations and translations), leaving invariant a distinguished direction which is normal to a predominant plane in the scene (e.g. a table used as support for the manipulations).

This article is structured as follows: in the next section we present preliminary concepts used in the paper, such as the Gauss Linking Integral, and we explain its limitations in a planar setting. Then, in Section 3 we introduce the novel concept of the directional derivative of the *GLI* which is also applicable for flat configurations. We derive first an expression for the *GLI* of two segments, then we prove that we can perturb the segments slightly to obtain information when they are co-planar and we explain how to apply this to a full meshed cloth. Next we study some of the properties of this new index by applying it to a database of cloth configurations of a napkin taken from simulated folding sequences, to then experimentally test the index on real images. We also show that it is possible to apply our representation to garments with non-trivial topology, such as a pair of shorts. Finally, we discuss the obtained results and draw some conclusions in the last section.

2. Preliminaries and related work

Definition 1 (Gauss Linking Integral of two curves). Given two \mathcal{C}^2 non-intersecting 3D-space curves γ_1, γ_2 parameterized by $x(s)$ and $y(t)$, respectively, with $s, t \in I = [0, 1]$, the *Gauss Linking Integral* between them, *GLI*

for short, is

$$GLI(\gamma_1, \gamma_2) = \frac{1}{4\pi} \int_I \int_I \frac{(y(t) - x(s)) \cdot (y'(t) \times x'(s))}{|y(t) - x(s)|^3} dt ds$$

or written in a compact way

$$GLI(\gamma_1, \gamma_2) = \frac{1}{4\pi} \int \int \frac{(\gamma_2 - \gamma_1) \cdot [\gamma_2' \times \gamma_1']}{\|\gamma_2 - \gamma_1\|^3}. \quad (1)$$

This double integral is invariant under re-parameterizations of the curves. In the case that both curves γ_1 and γ_2 are closed and smooth, their GLI is integer valued (due to the chosen normalization factor $\frac{1}{4\pi}$) and it is an invariant of the topology of the embedded curves (their *linking number*, see [1]).

Historically, the GLI was first introduced by Gauss, presumably related to his works on magnetism (according to [27]) or on astronomy (according to [11]). Considering the $GLI(\gamma, \gamma)$ of twice the same non-self-intersecting smooth curve γ , then the double integral (taking the domain of integration outside the diagonal of $I \times I$) defines another geometric invariant of the curve, known as *writhe* or *writhing number* of γ . Despite their resemblance, the GLI and the writhe measure different quantities: consider a normal vector field v of length $\epsilon > 0$ on γ , and the curve γ_v of endpoints of the vector field v , which is embedded and in one-to-one smooth correspondence with γ for sufficiently small ϵ . Then the GLI of these two close copies of the same γ differs from the writhe in $GLI(\gamma, \gamma_v) - GLI(\gamma, \gamma)$ equal to the total twist of v . This result is known as the Călugăreanu-White-Fuller theorem (see [23]). However, both indexes, GLI and writhe, are non-informative for planar curves, since they both vanish.

The GLI has been used for many applications after a version of the above formula for polygonal curves appeared in the context of DNA protein structures [19], with additional efficient formulations given in [18] from which we have chosen the following: given two piece-wise linear curves of N and M segments, that is, $\gamma_1 = \{\gamma_{P_i P_{i+1}}, i = 1, \dots, N\}$ and $\gamma_2 = \{\gamma_{Q_i Q_{i+1}}, i = 1, \dots, M\}$, where each segment is parameterized as $\gamma_{AB}(s) = A + s\vec{AB}$ for $s \in [0, 1]$, then the GLI between both curves is

$$GLI(\gamma_1, \gamma_2) = \frac{1}{4\pi} \sum_{i=1}^N \sum_{j=1}^M GLI(\gamma_{P_i P_{i+1}}, \gamma_{Q_j Q_{j+1}}) \quad (2)$$

where the *GLI* between a pair of segments γ_{AB} and γ_{CD} is computed as

$$\begin{aligned} GLI(\gamma_{AB}, \gamma_{CD}) = & \arcsin(\vec{n}_A \vec{n}_D) + \arcsin(\vec{n}_D \vec{n}_B) \\ & + \arcsin(\vec{n}_B \vec{n}_C) + \arcsin(\vec{n}_C \vec{n}_A) \end{aligned} \quad (3)$$

with

$$\begin{aligned} \vec{n}_A = & \|\vec{AC} \times \vec{AD}\|, \quad \vec{n}_B = \|\vec{BD} \times \vec{BC}\|, \\ \vec{n}_C = & \|\vec{BC} \times \vec{AC}\|, \quad \text{and } \vec{n}_D = \|\vec{AD} \times \vec{BD}\|. \end{aligned}$$

Remark 2.1. The above formula is *not* an approximation in the sense that it is the exact value of the integral (1) when applied to piece-wise linear curves.

The discrete formula (3) was used by Ho [15] to identify and synthesize animated characters in intertwined positions [14, 15]. In the context of robotics, the *GLI* has been applied to representative curves of the workspace to guide path planning through holes [16, 37], for guiding caging grasps in [24, 30, 31], and for planning humanoid robot motions using the *GLI* to guide reinforcement learning [36]. In this work, for the first time, we develop a further analysis of the notion to be able to apply it to planar or almost planar curves, which opens the door to a wider spectrum of applications.

3. Derivation of the Cloth Coordinates

As we have mentioned above, the *GLI* of two coplanar curves vanishes; so for many configurations of robotic interest — configurations where the cloth is nearly flat on a table, ready to be folded or already folded— the *GLI* does not provide much information. Our aim in this section is therefore to develop a similar index which is able to distinguish planar configurations. We shall see that a natural index to consider is in fact a directional derivative of the *GLI*, but to arrive at such an index we must first make a few observations about the *GLI* when applied to pairs of segments as in Equation (3), since the class of curves we will be working with computationally are piece-wise linear.

3.1. *GLI* of two segments

Since two segments AB and CD are uniquely defined by the four endpoints $A, B, C, D \in \mathbb{R}^3$, the *GLI* of two segments computed in Equation (3)

can be viewed as a function from $(\mathbb{R}^3)^4 \equiv \mathbb{R}^{12}$ to \mathbb{R} . To emphasize that from now on we are considering segments we define $\mathcal{G} : \mathbb{R}^{12} \rightarrow \mathbb{R}$ as

$$\mathcal{G}(A, B, C, D) = GLI(\gamma_{AB}, \gamma_{CD}) = \frac{1}{4\pi} \int \int \frac{(\gamma_{CD} - \gamma_{AB}) \cdot [\gamma'_{CD} \times \gamma'_{AB}]}{\|\gamma_{CD} - \gamma_{AB}\|^3}. \quad (4)$$

Note that technically \mathcal{G} is not defined in the whole of \mathbb{R}^{12} , since it is not defined when γ_{AB} and γ_{CD} intersect.

Lemma 1. Given two non-intersecting segments AB and CD with endpoints $A, B, C, D \in \mathbb{R}^3$, let γ_{CD} and γ_{AB} denote a parametrization of them. Then

$$\mathcal{G}(A, B, C, D) = \mathcal{V}(A, B, C, D) \cdot \mathcal{I}(A, B, C, D)$$

where $\mathcal{V}(A, B, C, D) = \det(\vec{AB}, \vec{AC}, \vec{AD})$ is the signed volume of the tetrahedron $ABCD$ multiplied by 6 and

$$\mathcal{I}(A, B, C, D) = \frac{1}{4\pi} \int \int \frac{1}{\|\gamma_{CD} - \gamma_{AB}\|^3}.$$

Proof. This is a straightforward computation: notice that the numerator in the integral expression of the GLI in Equation (1) is constant (for any t and s) because the curves are segments and equals

$$\begin{aligned} (\gamma_{CD} - \gamma_{AB}) \cdot [\gamma'_{CD} \times \gamma'_{AB}] &= \\ &= (\vec{AC} + t\vec{CD} - s\vec{AB}) \cdot [\vec{CD} \times \vec{AB}] = \\ &= \vec{AC} \cdot [\vec{CD} \times \vec{AB}] = \\ &= \vec{AC} \cdot [(\vec{CA} + \vec{AD}) \times \vec{AB}] = \\ &= \vec{AC} \cdot [\vec{AD} \times \vec{AB}] = \vec{AB} \cdot [\vec{AC} \times \vec{AD}] = \\ &= \det(\vec{AB}, \vec{AC}, \vec{AD}) \end{aligned}$$

the signed volume of the tetrahedron $ABCD$ multiplied by 6. By writing

$$\mathcal{V}(A, B, C, D) = \det(\vec{AB}, \vec{AC}, \vec{AD})$$

and

$$\mathcal{I}(A, B, C, D) = \frac{1}{4\pi} \int \int \frac{1}{\|\gamma_{CD} - \gamma_{AB}\|^3},$$

we have

$$\mathcal{G} = \mathcal{V} \cdot \mathcal{I}. \quad (5)$$

□

Remark 3.1. When two segments are co-planar their GLI vanishes since the tetrahedron they span has volume 0.

Corollary 3.1. The function $\mathcal{G}(A, B, C, D) = GLI(\gamma_{AB}, \gamma_{CD})$ is the product of two differentiable functions and hence differentiable with respect to A, B, C, D .

3.2. Directional derivative of \mathcal{G}

In this section we discuss how to perturb \mathcal{G} in order to make it informative in planar settings.

Definition 2 (Directional derivative of \mathcal{G}). Let $v_A, v_B, v_C, v_D \in \mathbb{R}^3$ be arbitrary directions and AB, CD two non-intersecting segments. The directional derivative of \mathcal{G} at the point (A, B, C, D) in the direction of $v = (v_A, v_B, v_C, v_D)$ is defined as the limit

$$\partial_v \mathcal{G}(A, B, C, D) = \lim_{\varepsilon \rightarrow 0} \frac{\mathcal{G}((A, B, C, D) + \varepsilon(v_A, v_B, v_C, v_D)) - \mathcal{G}(A, B, C, D)}{\varepsilon}.$$

Remark 3.2. Notice that this limit always exists since we have shown that \mathcal{G} is a differentiable function with respect to A, B, C, D . Moreover, $\partial_v \mathcal{G}$ can be equivalently written as

$$\lim_{\varepsilon \rightarrow 0} \frac{GLI(\gamma_{A^*B^*}, \gamma_{C^*D^*}) - GLI(\gamma_{AB}, \gamma_{CD})}{\varepsilon},$$

where $A^* = A + \varepsilon v_A$, $B^* = B + \varepsilon v_B$, $C^* = C + \varepsilon v_C$, $D^* = D + \varepsilon v_D$ and ε is sufficiently small so that A^*B^* and C^*D^* do not intersect. Also, from Equation (5) and by the product rule

$$\partial_v \mathcal{G} = \partial_v(\mathcal{V})\mathcal{I} + \mathcal{V}\partial_v(\mathcal{I}),$$

hence $\partial_v \mathcal{G} = \partial_v(\mathcal{V})\mathcal{I}$ when the segments γ_{AB} and γ_{CD} are coplanar.

Properties of $\partial_v \mathcal{G}$. By definition $\partial_v \mathcal{G}$ is invariant under translations, rotations and scalings if v is rotated and scaled accordingly. These properties are a consequence of the fact that the GLI is invariant under such transformations. However, for a fixed choice of v , $\partial_v \mathcal{G}$ will not be invariant under rotations or scalings in general. For instance, no fixed choice of v can make $\partial_v \mathcal{G}$ invariant under scalings, since scaling by a factor of λ scales \mathcal{I} by $\frac{1}{\lambda^3}$ and $\nabla \mathcal{V}$ by λ^2 , and similarly scales $\nabla \mathcal{I}$ by $\frac{1}{\lambda^4}$ and \mathcal{V} by λ^3 , resulting in scaling

$\nabla\mathcal{G}$ by $\frac{1}{\lambda}$. Depending on what this index is used for, one must keep this scaling relationship in mind or alternatively choose v depending on the segments. However, the distance we will use to compare different cloth states only depends on the correlation of values of the coordinates more than on the magnitude. That is why we can ignore the scaling factor that would appear when comparing two garments of different sizes (e.g. because of different meshings).

Choice of v . This is highly task-specific, but given the nature of our task—classifying planar cloth configurations based on affordances—it is natural to perturb the vertices in the direction normal to the table plane. Making such a choice of v does in fact make $\partial_v\mathcal{G}$ invariant under rotations and translations of the XY plane, which is desirable for our purposes since such movements of a cloth configuration have the same affordances. Furthermore, to conserve the symmetry $\partial_v\mathcal{G}(A, B, C, D) = \partial_v\mathcal{G}(C, D, A, B)$, we must perturb A and C by the same amount and direction, and the same is the case with B and D . Finally, it is easy to see that in fact perturbing A and C both by the same amount normal to the plane yields the same result as perturbing B and D by the same amount in the opposite direction, so it really only makes sense to perturb A and C , or B and D , but not both pairs, and doing one or the other is equivalent except for a sign change. In summary, the most natural choice of v in our case is

$$v := (\vec{0}, e_3, \vec{0}, e_3) \quad (6)$$

(or $v = (e_3, \vec{0}, e_3, \vec{0})$, which is equivalent except for a sign change) where $e_3 = (0, 0, 1)$ is the normal to the plane of the table on which the cloth lies.

3.3. Practical computation of $dGLI$

We summarize the discussion of the previous section in the following definition.

Definition 3 (dGLI of two segments). Given two non-intersecting segments γ_{AB} and γ_{CD} , we define

$$dGLI(\gamma_{AB}, \gamma_{CD}) := \lim_{\varepsilon \rightarrow 0} \frac{GLI(\gamma_{AB^*}, \gamma_{CD^*}) - GLI(\gamma_{AB}, \gamma_{CD})}{\varepsilon}, \quad (7)$$

where $B^* = B + \varepsilon e_3$, $D^* = D + \varepsilon e_3$, $e_3 = (0, 0, 1)$ and each GLI function can be computed using Equation (3). This index is invariant under rotations and translations of the XY plane and moreover $dGLI(\gamma_{CD}, \gamma_{AB}) = dGLI(\gamma_{AB}, \gamma_{CD})$.

Remark 3.3. We have analyzed numerically the limit defined in Equation (7), and have found that it is sufficiently stable to be computed as

$$dGLI(\gamma_{AB}, \gamma_{CD}) \cong \frac{GLI(\gamma_{AB*}, \gamma_{CD*}) - GLI(\gamma_{AB}, \gamma_{CD})}{\varepsilon}$$

for a sufficiently small ε . Since we work with double precision floats (which amount for a precision of around 14-15 decimals), it is known (see [12]) that when approximating derivatives numerically, one obtains better results when choosing perturbations which only affect 7 or 8 decimal places, for instance $\varepsilon \approx 10^{-8}$. This is the value taken in our experiments.

Remark 3.4. In practical implementations, we may well be computing the $dGLI$ between segments γ_{AB} and γ_{CD} that are very close to intersecting (but not intersecting since the cloth has thickness), and then $dGLI(\gamma_{AB}, \gamma_{CD})$ becomes very large. As having such big quantities can dominate values of metrics and distances in a non-representative way, in practice we set a maximum value to the $dGLI$ once it surpasses a fixed threshold.

3.4. Definition of the $dGLI$ Cloth Coordinates

Since we are now equipped with a geometric index for pairs of segments, we are ready to introduce our cloth coordinates, which will parametrize the shape-state space of a piece of cloth. We assign the $dGLI$ *Cloth Coordinates* to a cloth configuration \mathcal{C} of a garment as follows:

Definition 4 ($dGLI$ of a cloth surface \mathcal{C}). Given a discretization of the boundary of the garment surface \mathcal{C} as a polygonal curve, select an ordered subset of edges of it $\mathcal{S}_{\mathcal{C}} = \{S_i : i = 1, \dots, m\}$. Then, the $dGLI$ *Cloth Coordinates* of configuration \mathcal{C} is the upper triangular matrix

$$dGLI(\mathcal{C}) = (dGLI(S_i, S_j))_{S_i, S_j \in \mathcal{S}_{\mathcal{C}}, i > j}. \quad (8)$$

To get an intuitive sense of what these upper triangular matrices look like for some cloth configurations, see the examples in Figure 1. If we were interested in a general direction v , we would take the $dGLI_v$ *Cloth Coordinates*

$$dGLI_v(\mathcal{C}) = (\partial_v \mathcal{G}(S_i, S_j))_{S_i, S_j \in \mathcal{S}_{\mathcal{C}}, i > j}.$$

Note that the full matrix when taking all the edges of the discretization is the equivalent rationale to computing the GLI of a polynomial curve used

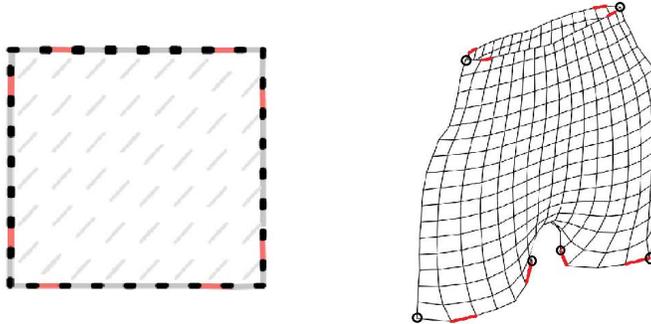


Figure 2: The subset of chosen segments are marked red. For the shorts, four of the selected boundary edges on the back are not visible in the figure.

in [15, 19], where the *GLI* of all pairs of segments of the curves were first assembled in what was there called the *GLI matrix* [15].

Choice of edges. The subset \mathcal{S}_c of edges chosen in the discretization depends on the task one wants to carry out; tasks that demand finer distinctions between configurations of a similar class would require a greater subset of segments. For our task of classifying the configurations into relatively broad classes, we found experimentally that a good choice of segments for a rectangular piece of cloth are the eight segments adjacent to the corner segments, marked red in the left panel of Figure 2. In the case of a garment with a more complicated topology, more edges must be chosen, e.g. for a pair of short pants we choose the twelve boundary edges shown in red in the right panel of Figure 2. In both cases, this small subset is nevertheless enough to provide an accurate affordance-based classification of near-to-flat configurations.

The upper triangular matrix in Equation (8), sorted as a vector, is a coordinate system that reduces the infinite dimensionality of the configuration space of cloth states to a mere $\frac{m(m-1)}{2}$ dimensional space. For instance, for a rectangular cloth $m = 8$, so this comes out to 28 dimensions. This reduction in dimensionality is well-suited and informative enough for practical purposes, as the validation results in next section will show.

4. Results

In this section we study the ability of the cloth coordinates previously defined to tell apart different cloth states. First, we analyze 4 folding sequences (see Figure 3). We will show that our representation is capable of distinguishing different relevant cloth configurations (e.g. one folded corner vs two folded corners). Then, we will apply our method to a full database with 12 cloth classes (shown in Figure 4), and we will compare it to 4 alternative representations, proving that ours is more capable of distinguishing between cloth states. All data in this section was simulated using the inextensible cloth model described in [7]. All simulations are performed with a square cloth of dimensions $1\text{m} \times 1\text{m}$ and a computational mesh of 400 nodes; except for the short pants (whose shape is taken from the UC Berkeley Garment Library) which consist of 537 nodes (see Figure 2, right). Finally, we will apply a simple classification method using our representation to real images of folded cloth states.

In order to compare different cloth configurations, once they are represented with our cloth coordinates $d\text{GLI}(\mathcal{C}) \in \mathbb{R}^{28}$, it is important to use a proper distance. Due to the scaling factor that we analyzed in the previous section, the most suitable distance was the *Spearman's distance*. Given two vectors x, y it is defined as

$$d(x, y) = 1 - \rho(\text{R}(x), \text{R}(y)), \quad (9)$$

where ρ is the Pearson correlation coefficient, and $\text{R}(x)$ is the rank variable of x (i.e. ordering the coordinates of x from lowest to greatest and then assigning to each coordinate its position in the ranking). This distance assesses how well the relationship between two vectors x, y can be described using any monotonic function (not only a line). We found this distance to be more sensitive to changes of the cloth configuration than the euclidean distance. This may be due to the fact that this distance focuses on the ranking order between coordinates (with sign) rather than comparing their magnitudes, which is most relevant in our representation. Note that the Spearman's distance is bounded with values between 0 and 2 and ignores scaling factors between different clothes.

4.1. Analysis of folding sequences

The first test compares different cloth states inside a folding sequence. Given the motion of the cloth $\{\mathcal{C}_1, \dots, \mathcal{C}_m\}$, where m is the number of discrete

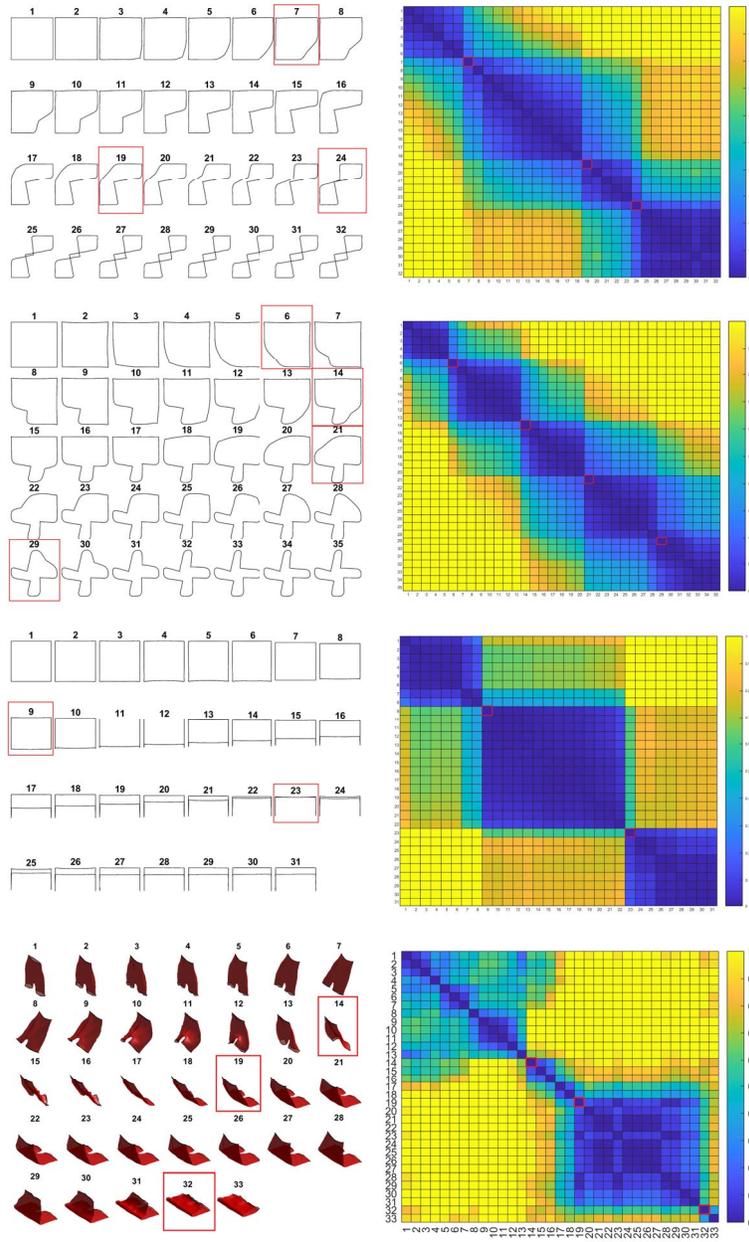


Figure 3: Study of the index during 4 folding sequences. In the left column we show a representation of the cloth frames, and in the right column the confusion matrix of all of them. In red we highlight the class changes that can be identified.

frames and \mathcal{C}_i is the state of the cloth at t_i , we compute the confusion matrix $\mathcal{M}_{ij} = d(\text{dGLI}(\mathcal{C}_i), \text{dGLI}(\mathcal{C}_j))$. The 4 folding sequences, shown at the left side of Figure 3 are: folding two opposite corners, folding 4 corners inwards, folding the cloth in half and folding a pair of shorts in half dynamically. The results can be seen on the right side of the figure. Notice how our representation detects changes during the sequence that are meaningful. For example, in Seq. 1, folding two opposite corners, at frame 7, there is an important change, since a corner changes the orientation from flat to folded, even before it is released. This can be seen in the confusion matrix (first two blue squares). This is also clear in Seq. 2, where four corners are folded inwards. Moreover, our method also detects when edges of the cloth cross (Seq. 1, frame 24, Seq. 3, frame 23). These changes are also meaningful from the manipulation point of view, as they afford different possible graspings or actions. Especial mention deserves Seq. 4 since the cloth has non-trivial topology and the garment hangs and is non-flat during a significant part of the folding. In this sequence our representation detects correctly the moment in which the shorts touch the table (frame 14), when the top two controlled corners start to descend (frame 19) and the moment in which the fold is completed (frame 32).

4.2. Confusion matrix of the full database

We now analyze a complete database consisting of 120 examples classified in 12 different classes of states, shown in Figure 4. Most of them are self-explanatory. Note that in the class 10 the upper left corner is folded under the cloth (likewise for class 11). Each class has 10 samples corresponding to the final state of the cloth during a folding sequence simulation. We manually identified samples that we considered to belong to the same state. We want to emphasize that once we fix an ordering of the corners, our method distinguishes, for example, between different folded corners and this does not contradict the rotational invariance previously shown.

Again we compute the confusion matrix $\mathcal{M}_{ij} = d(\text{dGLI}(\mathcal{C}_i), \text{dGLI}(\mathcal{C}_j))$ where \mathcal{C}_k is the k th example of the database. We order the samples, so that the samples from the same classes are consecutive. This way, the plot is more easily interpretable. In Figure 4 we can see how the classes group without confusion: i.e. the distance between members of a class tends to be smaller (color blue) than the distance to examples outside the class (color yellow).

The confusion matrix shows us interesting insights about our representation. For instance, we can see that the two classes 01 and 04 are relatively

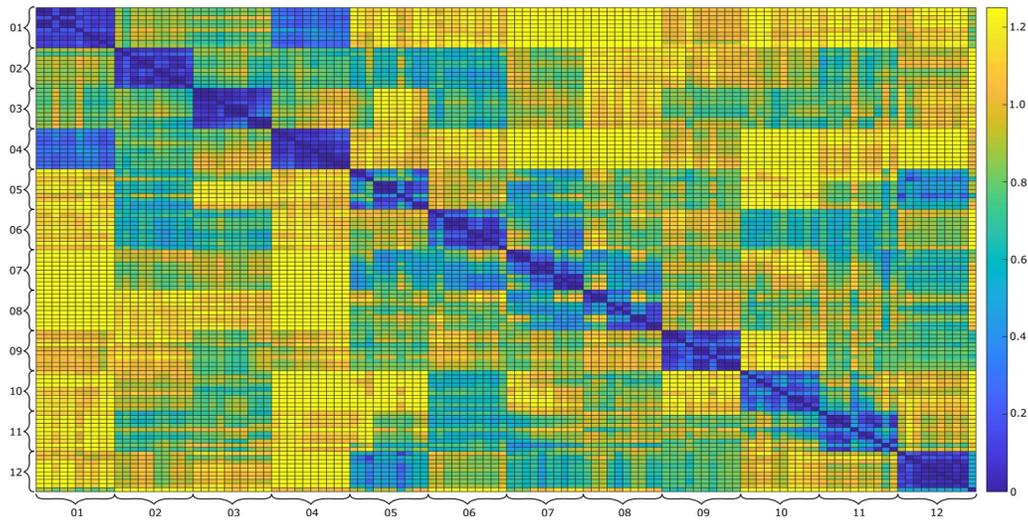
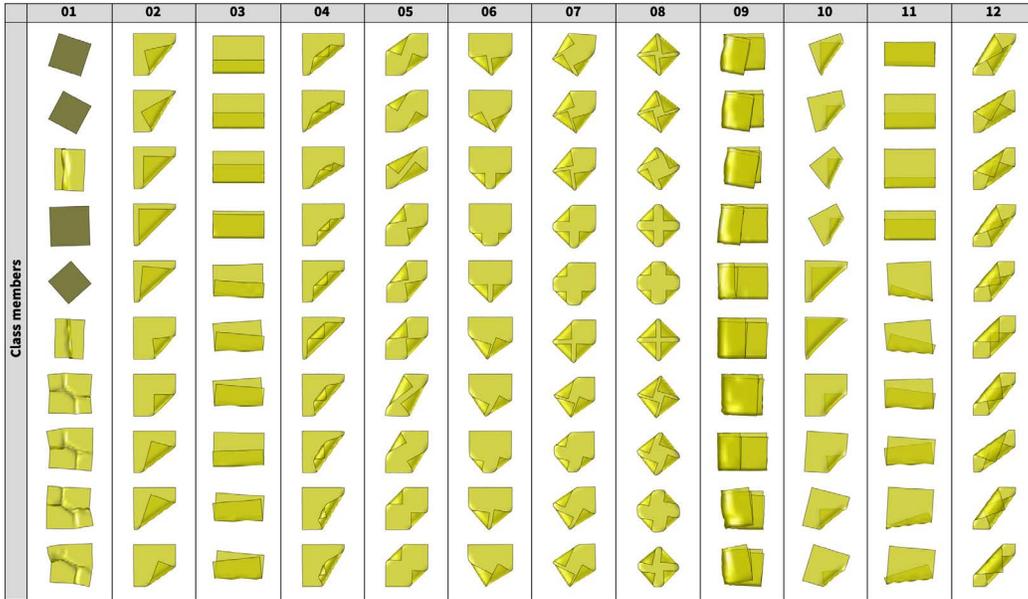


Figure 4: Confusion matrix of all the distances between the states shown in the top table.

closer than others. That is because the relative position of all edges is indeed the same in these classes, resulting in a smaller distance in our representation. The same phenomenon can be seen between classes 05 and 12 in some cases, as they are indeed classes with similarities (in 05 the two corners do

not cross, whereas in 12 they do). However, classes 03 and 11, which differ on whether or not the folding makes one side of the cloth hide its opposite, are perfectly separated. The borderline cases, that is, the fourth element in class 03 and the first element in class 11 are very similar, but our method distinguishes them because of the relative geometric position between edges (i.e. in these two cases, they are flipped). A similar thing occurs between classes 02 and 10. It is also worth mentioning that some classes that we have labeled as the same class have clear sub-classes shown in the confusion matrix. That is the case for classes 05, 07 and 08. These are folded corners with different orientations. It is possible, using our representation, to induce a partition of the space in order to separate this type of class into two.

4.3. Comparison with other shape representations

In this subsection we perform a more quantitative comparison of our state representation with other competing methods in representing shapes. To evaluate a representation, we use the standard Davies-Bouldin index to measure cluster separation [8]:

$$DB = \frac{1}{n} \sum_{i=1}^n \max_{j \neq i} \left(\frac{\sigma_i + \sigma_j}{d(c_i, c_j)} \right) \quad (10)$$

where n is the number of classes (e.g. in the database is 12), c_i is the centroid of class i (the average of the coordinates of members of class i), σ_i is a dispersion measure computed as the average distance of all elements in class i to the centroid c_i and $d(c_i, c_j)$ is the distance between centroids c_i and c_j . With the classification given in Figures 3 and 4 taken as ground truth, we want a representation that gives a small dispersion inside a class and high distance between the classes, resulting in a low index. The representation and distance with the smallest DB is considered the one that better separates these clusters, and therefore, the best representation to identify different cloth states.

First, we use two simple cloth shape representations using similar low-level features like the ones we used:

- (i) **Edges:** for a given mesh we select the edges shown in Figure 2 and compute their pairwise minimal distance. This results in a representation vector of length just like that of the dGLI coordinates (notice that unlike the dGLI, the coordinates of this vector are always non-negative). We use the Spearman’s distance to compare two different

Table 1: Comparison between different shape representations*

Database	Seq. I	Seq. II	Seq. III	Seq. IV	
dGLI	0.73	0.27	0.18	0.21	0.52
Edges	1.60	0.68	0.77	0.51	1.91
Corners	2.49	0.98	1.61	3.14	1.86
Fréchet	0.99	0.69	0.76	0.48	0.90
Hausdorff	1.45	0.71	0.84	0.49	0.90

*Each number is the Davies-Bouldin index introduced in Equation 10, that measures cluster separation quality. A smaller value means a better separation. We mark in bold the smallest values in each column.

samples. This representation is invariant under rigid motions of the plane.

- (ii) **Corners:** for a given rectangular mesh we compute the pairwise distance between its 4 corners. In the case of the shorts, we take the 6 nodes shown in Figure 2. These are 6 or 15 non-negative numbers that can be computed for any cloth, they are invariant by rigid motions and they give a trivial representation of the state of the cloth. We also use Spearman’s distance to compare different samples.

In addition, we compare with two classic methods to measure distance between curves and polygons [4, 33], taking the full discrete boundary curve of the garments as the state representation:

- (iii) **Fréchet:** to compare two different samples we compute the (discrete) Fréchet distance [10] between the curves. This is a distance that takes into account the location and ordering of the points along the curves. Since this distance is not invariant by rigid motions, special care must be taken to center and align the samples before comparing them. In order to do so we center the curves at the origin and perform a rigid alignment by computing the rotation that minimizes the distance between the curves’ points.
- (iv) **Hausdorff:** to compare two different samples we compute the (discrete) Hausdorff distance between the points of the curves [13]. Informally, two curves are close in the Hausdorff sense if every point of

either curve is close to some other point of the other one. This distance disregards the fact that the sets it is comparing are curves and therefore is expected to be less sensitive than the Fréchet distance. As before, since this distance is not invariant by rigid motions, we center and align the samples before comparing them.

In Table 1, we display the computation of the DB index for our dGLI coordinates and the four discussed methods, using as testing scenarios the full database and the 4 folding sequences presented before. As seen in the table, our method results in the lowest overall DB in all 5 scenarios, indicating that our method is the one among those studied that best represents the different folded states of the textiles.

4.4. Real images classification

Once checked that our method was able to represent folded states of cloth accurately, we implemented a simple classifier of real folded cloth states to assess its applicability. In order to do so, a synthetic representative element of each class in the database shown in Figure 4 is chosen, and we estimate the class of a new real unclassified sample by choosing its closest representative, using Spearman’s distance.

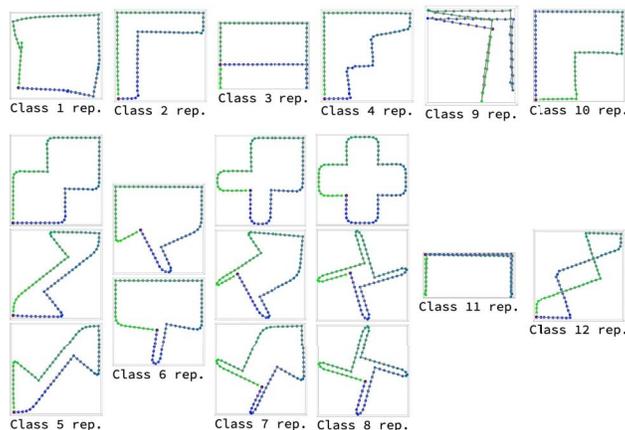


Figure 5: Synthetic representatives chosen for each class. When only one is chosen, it is the closest to the centroid of the class. When a class has more sparsity, additional representatives are chosen to represent the subgroups in the class.

The real images are taken from a zenital position at 52 cm from the table using a Microsoft Azure Kinect DK 3D camera. A single napkin is

used with 3 colored stickers attached along each edge, close to a corner and on both sides. We first use color segmentation to detect the center of each sticker and get the corresponding 3D point from the depth image. Once all markers are detected, with our combinations of colors on each edge, we can identify each individual corner of the cloth (there are four stickers of the same color around each corner), and its corresponding edge positions, following the same edge selection as in Figure 2. The obtained size of the observed edges is more than 400 times larger than the edges of the samples of the simulated database, but thanks to the Spearman’s distance used, this does not affect the distance values when comparing shapes of different sizes.

As we can see in the confusion matrix in Figure 4, some classes have a larger dispersion in distance because of the variation in orientations of the corners. For these classes, we have chosen 3 different representatives, corresponding to the three subgroups that can be clearly seen in the confusion matrix. We show the silhouette of the representatives chosen for each class in Figure 5. The table in Figure 6 shows the results of the classification.

The only miss-classification is that of the last image in class 04. However, note that this is a very extreme case where the edges of the cloth are in a relative position very similar to the flat unfolded case, and therefore, it is classified in class 01. This is a reasonable mistake, as this cloth can be considered flat enough.

Notice that we can only perceive those textiles with all the stickers visible, therefore, classes with hidden edges, like for instance classes from 09 to 12 where the folding is under the cloth, are not present in the real set of samples. However, the classifier still used all 12 classes of the simulated database. This shows that the missing classes don’t create confusion in the classification process.

5. Conclusions

We have proposed the dGLI *Cloth Coordinates*, a representation for cloth configurations based on a directional derivative of a topological index that greatly reduces the dimensionality of the cloth configuration space, going from a full cloth surface to a vector of dimension 28 (for a rectangular cloth). This reduced representation nevertheless preserves enough information about the configurations to be able to distinguish them according to their grasping affordances using Spearman’s distance. The fact that using our representation we can successfully classify real configurations of cloth from synthetic

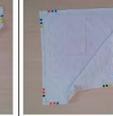
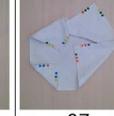
GT	Images of cloth with predicted class						
02	 02	 02	 02	 02	 02	 02	 02
03	 03	 03	 03	 03	 03	 03	
04	 04	 04	 04	 04	 04	 04	 01
05	 05	 05	 05	 05	 05	 05	 05
06	 06	 06	 06	 06			
07	 07	 07	 07	 07	 07	 07	
08	 08						

Figure 6: Results of the real image classification using the simulated database presented in Figure 4 as reference. The first column shows the ground truth class of the images, and at the bottom of every image the classified class.

generated samples as seen in Figure 6 shows great promise for applications in planning for cloth manipulation. Furthermore, our representation allows for different choices of v , the perturbation direction, and \mathcal{S} , the subset of edges chosen, so that one can fine-tune the representation to the specific task at hand to boost results. Moreover, we evaluated successfully the ex-

pressive power of our representation during a folding sequence of a garment with non-trivial topology (a pair of shorts). Lastly, since our method is not learning-based, it does not require any training data, it is completely explainable, and it is robust against possible configurations that are not in the training set.

In summary, the dGLI Cloth Coordinates bridge the gap between low-level features of different cloth configurations, such as the location of corners and edges, to high-level semantic identification of cloth states, associated to their possible affordances.

6. Limitations and further work

Although a strong assumption is made in this work, that is, that we know the full border of the cloth, perception algorithms are starting to show solutions to overcome this problem. For instance, in [26] a method is developed to detect parts of clothes suitable for grasping and more recently, the deep-learning approach presented in [25] can identify corners and edges, but does not yet identify the full border. Our group is working on different deep-learning and mathematical methods to *hallucinate* the full boundary given an image (or point-cloud) of the cloth to overcome this limitation.

Meanwhile, our representation can be fully used in simulation with several important applications, such as building datasets where automatic segmentation of the cloth states is required (see e.g. [5]), monitor cloth manipulation and guide planning methods. We are looking forward to pursue all these lines of research that the present work opens the door to.

Future work also concerns an in-depth analysis of the configuration space defined by our coordinates. In particular, we would like to identify a partition of the space that corresponds to a partition of configurations by grasping affordance, which states are neighbors in this partition, and what the shortest paths from one state to another are. This line of research is expected to be especially complex and rich for garments with non-trivial topology. We look forward to carrying out this study analytically as well as through learning methods, which we believe will give better results when the data is enriched and given structure through our representation.

7. Acknowledgments

This work was developed under the project CLOTHILDE which has received funding from the European Research Council (ERC) under the

EU-Horizon 2020 research and innovation programme (grant agreement No. 741930). M. Alberich-Carramiñana is also with the Barcelona Graduate School of Mathematics (BGSMath) and the Institut de Matemàtiques de la UPC-BarcelonaTech (IMTech), and she and J. Amorós are partially supported by the Spanish State Research Agency AEI/10.13039/501100011033 grant PID2019-103849GB-I00 and by the AGAUR project 2021 SGR 00603 Geometry of Manifolds and Applications, GEOMVAP. J. Borràs is supported by the Spanish State Research Agency MCIN/ AEI /10.13039/501100011033 grant PID2020-118649RB-I00 (CHLOE-GRAPH project).

References

- [1] J. Adlinger, I. Klapper, and M. Tabor. Formulae for the calculation and estimation of writhe. *Journal of Knot Theory and Its Ramifications*, 4:343–372, 1995.
- [2] E. E. Aksoy, A. Abramov, J. Dörr, K. Ning, B. Dellen, and F. Wörgötter. Learning the semantics of object–action relations by observation. *International Journal of Robotics Research.*, 30(10):1229–1249, 2011.
- [3] M. Alberich-Carramiñana, J. Amorós, and F. Coltraro. *Developable surfaces with prescribed boundary*. In Extended Abstracts GEOMVAP 2019, Springer-Birkhaeuser., 2021.
- [4] H. Alt and M. Godau. Computing the fréchet distance between two polygonal curves. *International Journal of Computational Geometry & Applications*, 5(01n02):75–91, 1995.
- [5] 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, international conference on robotics and automation. In *To appear on 2023 IEEE International Conference on Robotics and Automation (ICRA 2023)*. IEEE, 2023.
- [6] C. Bousquet-Jette, S. Achiche, D. Beaini, Y. Law-Kam Cio, C. Leblond-Ménard, and M. Raison. Fast scene analysis using vision and artificial intelligence for object prehension by an assistive robot. *Engineering Applications of Artificial Intelligence*, 63:33–44, 2017.

- [7] F. Coltraro, J. Amorós, M. Alberich-Carramiñana, and C. Torras. An inextensible model for the robotic manipulation of textiles. *Applied Mathematical Modelling*, 101:832–858, 2022.
- [8] D. L. Davies and D. W. Bouldin. A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-1(2):224–227, 1979.
- [9] A. Doumanoglou, J. Stria, G. Peleka, I. Mariolis, V. Petrik, A. Kargakos, L. Wagner, V. Hlavac, T.-K. Kim, and S. Malassiotis. Folding clothes autonomously: A complete pipeline. *IEEE Transactions on Robotics*, 32(6):1461–1478, 2016.
- [10] T. Eiter and H. Mannila. Computing discrete fréchet distance. May 1994.
- [11] M. Epple. Orbits of asteroids, a braid, and the first link invariant. *The Mathematical Intelligencer*, 20:45–52, 1998.
- [12] J. Faires and R. Burden. *Numerical Methods, 4th*. Cengage Learning, 2012.
- [13] J. Henrikson. Completeness and total boundedness of the hausdorff metric. *MIT Undergraduate Journal of Mathematics*, 1(69-80):10, 1999.
- [14] E. S. Ho, T. Komura, S. Ramamoorthy, and S. Vijayakumar. Controlling humanoid robots in topology coordinates. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 178–182. IEEE, 2010.
- [15] S. L. Ho. *Topology-based character motion synthesis*. PhD thesis, University of Edinburgh, 2011.
- [16] V. Ivan, D. Zarubin, M. Toussaint, T. Komura, and S. Vijayakumar. Topology-based representations for motion planning and generalization in dynamic environments with interactions. *The International Journal of Robotics Research*, 32(9-10):1151–1163, 2013.
- [17] R. Jangir, G. Alenya, and C. Torras. Dynamic cloth manipulation with deep reinforcement learning. In *IEEE International Conference on Robotics and Automation*, pages 4630–4636, 2020.

- [18] K. Klenin and J. Langowski. Computation of writhe in modeling of supercoiled DNA. *Biopolymers*, 54(5):307–317, 2000.
- [19] M. Levitt. Protein folding by restrained energy minimization and molecular dynamics. *Journal of molecular biology*, 170(3):723–764, 1983.
- [20] M. Lippi, P. Poklucar, M. C. Welle, A. Varava, H. Yin, A. Marino, and D. Kragic. Latent space roadmap for visual action planning of deformable and rigid object manipulation. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5619–5626, 2020.
- [21] J. Matas, S. James, and A. J. Davison. Sim-to-real reinforcement learning for deformable object manipulation. In *Proceedings of Conference on Robotic Learning*, 2018.
- [22] S. Miller, J. Van Den Berg, M. Fritz, T. Darrell, K. Goldberg, and P. Abbeel. A geometric approach to robotic laundry folding. *The International Journal of Robotics Research*, 31(2):249–267, 2012.
- [23] W. F. Pohl. DNA and differential geometry. *The Mathematical Intelligencer*, 3:20–27, 1980.
- [24] 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*, pages 1100–1107. IEEE, 2013.
- [25] J. Qian, T. Weng, L. Zhang, B. Okorn, and D. Held. Cloth region segmentation for robust grasp selection. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 9553–9560, 2020.
- [26] A. Ramisa, G. Alenyà, F. Moreno-Noguer, and C. Torras. A 3d descriptor to detect task-oriented grasping points in clothing. *Pattern Recognition*, 60:936–948, 2016.
- [27] R. L. Ricca and B. Nipoti. Gauss linking number revisited. *Journal of Knot Theory and Its Ramifications*, 20:1325–1343, 2011.
- [28] 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. *The International Journal of Robotics Research*, 37(7):688–716, 2018.

- [29] D. Seita, A. Ganapathi, R. Hoque, M. Hwang, E. Cen, A. K. Tanwani, A. Balakrishna, B. Thananjeyan, J. Ichnowski, N. Jamali, et al. Deep imitation learning of sequential fabric smoothing policies. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 9651–9658, 2020.
- [30] J. A. Stork, F. T. Pokorny, and D. Kragic. Integrated motion and clasp planning with virtual linking. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3007–3014. IEEE, 2013.
- [31] J. A. Stork, F. T. Pokorny, and D. Kragic. A topology-based object representation for clasping, latching and hooking. In *2013 13th IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, pages 138–145. IEEE, 2013.
- [32] D. Tanaka, S. Arnold, and K. Yamazaki. Emd net: An encode–manipulate–decode network for cloth manipulation. *IEEE Robotics and Automation Letters*, 3(3):1771–1778, 2018.
- [33] R. C. Veltkamp and M. Hagedoorn. State of the art in shape matching. *Principles of visual information retrieval*, pages 87–119, 2001.
- [34] F. Wörgötter, E. E. Aksoy, N. Krüger, J. Piater, A. Ude, and M. Tamosiunaite. A simple ontology of manipulation actions based on hand-object relations. *IEEE Transactions on Autonomous Mental Development*, 5(2):117–134, 2013.
- [35] H. Yin, A. Varava, and D. Kragic. Modeling, learning, perception, and control methods for deformable object manipulation. *Science Robotics*, 6(54):eabd8803, 2021.
- [36] W. Yuan, K. Hang, H. Song, D. Kragic, M. Y. Wang, and J. A. Stork. Reinforcement learning in topology-based representation for human body movement with whole arm manipulation. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 2153–2160. IEEE, 2019.
- [37] 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.