

Active Boundary Component Models for Robotic Dressing Assistance

Lukas Twardon and Helge Ritter

Abstract—The dynamics of deformable objects, especially that of highly flexible articles of clothing, is difficult to model. This is due to their vast number of degrees of freedom in addition to the noisy and incomplete measurements robots have to cope with. Therefore, we suggest focusing on the structures and object parts which are relevant to the task at hand. The openings (e.g., at the waist, leg or sleeve ends) characterize garments surprisingly well, not only from a topological perspective, but also in terms of their inherent function, namely dressing. We model openings as closed, oriented chains of movable points which we refer to as *Active Boundary Component Models (ABCMs)*. Compared with the hardly predictable motions of an overall piece of clothing, relatively strict assumptions regarding the dynamics of these contour models can be made. We express these assumptions through position-based constraints which drastically restrict the degrees of freedom. In the present paper, we show how *ABCMs* can be initialized exploiting geometric prior knowledge of garments, and how they can be tracked visually using 3D point cloud data. Additionally, we consider the task of sliding a rod through a pant leg as a first step toward robotic dressing assistance for physically handicapped persons.

I. INTRODUCTION

Robots are making good progress in their ability to perceive and manipulate all kinds of everyday objects. While the dynamics of rigid bodies is comparatively well understood, it is only in recent years that roboticists have begun to focus on objects with high-dimensional configuration spaces. The extreme deformability of most articles of clothing hence poses a complex challenge to present-day robots. Nevertheless, clothing plays an important role in people’s everyday lives, and the ability to handle clothes reliably would make robots much more practical for home use.

Researchers have studied clothing classification [1], pose estimation [2]–[5], grasping [6], folding [7]–[10], and unfolding [11]–[13], as well as flattening [14]. One task that is worth taking a closer look at is robotic dressing assistance [15]–[17]. This is, on the one hand, because it could be exceptionally useful for handicapped persons, and on the other hand, because putting them on is the inherent function of all garments. This function is reflected in their designs and affordances, and we should keep it in mind when trying to model clothing computationally.

There are two general ways of how clothes can be represented geometrically. In many cases, it is sufficient to consider a 2D or 2.5D projection on a plane which is very often the tabletop. In other cases, a full 3D representation of the object or parts of the object is needed. The choice mainly

depends on the task to be performed. Most existing works on robotic manipulation of clothes clearly fall into one of the two categories.

Both Miller et al. [7] and Stria et al. [8] initialized 2D polygonal models from clothes roughly spread out on a flat surface and used these models to implement a folding procedure. The robot in [14] performed a sequence of flattening actions based on a surface analysis of a 2.5D heightmap. Ramisa et al. [6] used a similar representation to define a measure of wrinkledness for optimal grasp pose estimation.

Unfolding a randomly hanging garment is a complex task that requires a detailed understanding of the 3D structure and pose of the object. Doumanoglou et al. [11] approached the unfolding problem with some reasoning about possible lowest points together with random forest based recognition and probabilistic planning. Li et al. [2] proposed a two stage recognition scheme consisting of an offline simulation phase and an online matching and registration phase. They used their method for both robotic unfolding [12] and folding [9] of a piece of clothing. Bersch et al. [10] printed fiducial markers on clothes in order to simplify recognition, whereas Willimon et al. [3] suggested a markerless approach to 3D pose estimation of deformable surfaces. Furthermore, Willimon et al. applied the interactive perception paradigm to the tasks of unfolding [13] and classification [1] of clothing. A method employing physics-based simulation that produced some impressive deformable object tracking results was described in [4]. Kita et al. [5] showed that strategic observation can be helpful for 3D shape estimation of garments.

To date, only very few works have considered the dressing or dressing assistance problem. Colomé et al. [15] addressed the task of wrapping a scarf around the neck of a mannequin. While their focus was on safety in human-robot interaction, the actual task was much simpler than the general case because it did not involve garments with openings or holes. Tamei et al. [16], by contrast, explicitly modeled the relationship between a garment opening and a human body part through topology coordinates. Specifically, their robot learned to put a mannequin’s head into a T-shirt. In the initial setup, the T-shirt’s neck was equipped with markers for motion capturing, whereas Koganti et al. [17] reported some success in markerless estimation of the topology coordinates.

One of the main contributions of the present paper will be a more generic consideration of the problem. In Section II, we discuss the dressing task topologically, geometrically, and functionally which, to the best of our knowledge, has not been done before. Section III summarizes our previous work on robotic manipulation of clothes. The second key contribution will be a constraint-based approach to modeling

The authors are with the Neuroinformatics Group, EXC Cognitive Interaction Technology (CITEC), Bielefeld University, Germany. {ltwardon, helge}@techfak.uni-bielefeld.de

the dynamics of garment openings which we postulate to be essential for the proper handling of clothing. To this end, we elaborate the novel concept of *Active Boundary Component Models (ABCMs)* in Section IV. The spatial relationships between boundary components are represented through *Active Skeleton Models* (see Section V). Sections VI and VII describe how *ABCMs* can be initialized and tracked with point clouds, respectively. In Section VIII, we report the results of our experiments, including a preliminary dressing assistance test with a bimanual robot. Finally, we present our conclusions in Section IX.

II. PROBLEM STATEMENT

In the following, we formalize the task of putting on a piece of clothing by taking the examples of a sweater, a pair of pants, and a leg warmer (Fig. 1). The topology of a garment is defined by its openings which are represented by closed contours referred to as boundary components. Considering the exterior surface of an article of clothing as an orientable genus zero 2-manifold with boundary, the number of boundary components is the only remaining invariant up to homeomorphism. The considered garments have two (leg warmer), three (pair of pants), and four (sweater) boundary components, respectively¹. Geometrically, the boundary components are not necessarily ellipses, although they often are in their undeformed standard configuration. In general, the overall geometry of a garment can be very complex. However, when spread out on a flat surface, it usually takes a polygonal shape that is characteristic of its category. This 2D representation is related to the garment topology in that typically all boundary components appear as line segments of the polygon.

In order to specify the task of getting dressed, not only the article of clothing but also the human body has to be modeled. A very intuitive representation is that of a stick figure or skeleton. Commonly, only a sub-skeleton (such as the upper body, the lower body, or a single leg) is involved in putting on a garment. For simplicity, we assume a star-shaped sub-skeleton with a single central point. Dressing then begins with the identification of a specific boundary component b_0 (the upper end in case of a leg warmer, the waist end in case of a sweater or a pair of pants) using prior knowledge of the garment geometry². The pose of b_0 has to be tracked throughout the dressing process because all involved skeleton ends have to pass through this opening before sliding through the garment interior toward their respective target boundary components b_T . Suitable paths through the garment can be determined by considering the target skeleton configuration. An exemplary trajectory of a particular skeleton end then starts at b_0 , passes the central point of the target sub-skeleton,

¹The general consideration also holds for garments like button shirts where the collar forms a large boundary component together with the front and the waist part. However, optimal visual tracking of such contours may be different from the method described in this paper. Topology changes such as buttoning or zipping are not considered in the present work.

²Apart from some unusual dressing strategies which involve turning the item of clothing inside out, the choice of b_0 is uniquely predetermined by the garment category.

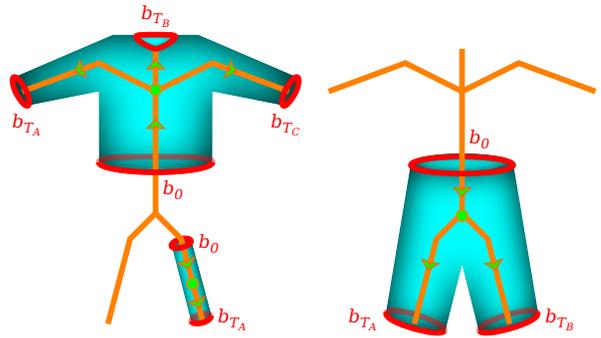


Fig. 1. Illustration of the dressing task, exemplified by a sweater, a pair of pants, and a leg warmer. The boundary components of the garments are shown in red. The orange-colored skeletons schematically represent human bodies. The central points of involved sub-skeletons and possible trajectories are depicted in green.

and ends at b_T . As is the case with boundary components, the trajectories may also deform when the garment is deformed. The task is completed as soon as each skeleton end matches its associated opening b_T .

III. PREVIOUS WORK

Many of the works on deformable object manipulation focus on textiles with trivial topologies such as table cloths or towels. Others are concerned with more complex clothes, but ignore their topological properties to the greatest possible extent. In a previous paper [18], we aimed to initiate a paradigm shift from detecting overly detailed metric features to identifying task-relevant topological structures which, in the case of garments, almost always involve openings. To this end, we developed an approach to extracting boundary components from point cloud data.

We employed a graph representation which was built through normal-based edge detection, skeletonization of the edge image, and contour following. Finding and evaluating simple cycles in the graph yielded the boundary components. We were able to define stable grasp poses based on the extracted boundary components, and to show their usefulness in a robotic coat-check scenario. However, the graph-based model has the drawback of being static and, hence, not allowing active tracking. Nevertheless, the technique is well suited for initializing *ABCMs*.

IV. ACTIVE BOUNDARY COMPONENT MODELS

As described in Section II, the boundary components of clothes play an important role in dressing assistance. Therefore, we introduce a framework for modeling the dynamics of these closed contours. We define *Active Boundary Component Models (ABCMs)* as tuples of 3D points with attached constraints (Fig. 2). The order in which the points are given specifies the direction of the corresponding opening in accordance with the right-hand rule. While classical contour tracking approaches such as *snakes* [19] or particle filter based methods [20] limit the degrees of freedom implicitly through energy terms or by choosing a certain curve parametrization, *ABCMs* provide an intuitive and extensible mechanism to

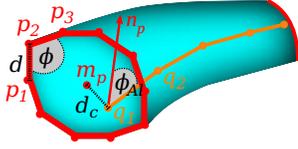


Fig. 2. Schematic diagram of an *ABCM* (red), an *Active Skeleton Model* (orange), and some of the constraint parameters (black).

express physical and topological prior knowledge explicitly in a constraint-based manner. Particularly, *ABCMs* have the following properties:

- *Simplicity*: As boundary components are simply approximated by a closed sequence of points, no complex curve parametrization is needed.
- *Flexibility*: *ABCMs* are in principle not restricted to visual tracking. Boundary points can be freely manipulated according to whatever forces or sensory modalities influence the object perception.
- *Plausibility*: Despite their flexibility, the models must behave plausibly in terms of the physical and topological assumptions made. The assumptions are formalized through constraints on the boundary points.
- *Stability*: We formulate the constraints in a position-based manner which ensures high stability and controllability as compared with most force-based methods.

At each update iteration, *ABCMs* undergo a two-step process. First, the positions of the boundary points are manipulated according to sensory input. Then, a constraint solving step ensures model plausibility. Position-based constraints are solved using constraint functions C and their gradients ∇C with respect to the involved points p_1, \dots, p_n as described by Müller et al. [21]. Let p be the concatenation vector $[p_1^T, \dots, p_n^T]^T$. Then, constraint solving means finding a correction Δp such that $C(p + \Delta p) = 0$ (equality constraint), $C(p + \Delta p) < 0$, or $C(p + \Delta p) > 0$ (inequality constraints). It can be shown [21] that the correction of an individual boundary point is given by

$$\Delta p_i = -c_s \frac{C(p)}{|\nabla_p C(p)|^2} \nabla_{p_i} C(p), \quad (1)$$

where $c_s \in [0, 1]$ determines the strength of the constraint. To solve multiple constraints, the corrections are applied repeatedly for all unsatisfied constraints, one after another.

One of the basic ideas of *ABCMs* is that even though the overall dynamics of clothes may be hard to predict, we can still make a few reasonable assumptions regarding the boundary components, in particular the following:

- 1) *Roughly constant arc length*: Boundary component shrinkage or expansion is minimal during typical garment manipulations.
- 2) *Smoothness*: Although different materials allow different degrees of deformation, model plausibility is drastically increased by assuming a minimum boundary smoothness.
- 3) *No entanglement*: The boundary components of real-world garments do not excessively coil out of the plane.

To formalize these assumptions, we now derive suitable constraint functions along with their gradients as required by Eq. (1).

Constant arc length is ensured by imposing equality constraints on the distances of adjacent boundary points p_1 and p_2 . The constraint function is

$$C_1(p_1, p_2) = |p_1 - p_2| - d, \quad (2)$$

where d is chosen to be the distance between the two points at initialization time. It is also possible to allow stretching by replacing the equality constraint with two inequality constraints indicating the lower and upper distance tolerances. The gradients are as in [21]:

$$\nabla_{p_1} C_1(p_1, p_2) = \frac{p_1 - p_2}{|p_1 - p_2|} \quad (3)$$

$$\nabla_{p_2} C_1(p_1, p_2) = -\frac{p_1 - p_2}{|p_1 - p_2|} \quad (4)$$

A natural way to model *smoothness* is through angular inequality constraints ($C_2(p_1, p_2, p_3) > 0$) between adjacent segments $\overline{p_1 p_2}$ and $\overline{p_2 p_3}$ of a boundary component, the constraint function being

$$C_2(p_1, p_2, p_3) = \arccos\left(\underbrace{\left(\frac{p_1 - p_2}{|p_1 - p_2|}\right)^T}_{\hat{a}} \underbrace{\left(\frac{p_3 - p_2}{|p_3 - p_2|}\right)}_{\hat{b}}\right) - \phi, \quad (5)$$

and the gradients with respect to the points being

$$\nabla_{p_1} C_2(p_1, p_2, p_3) = -\frac{1}{\sqrt{1 - (\hat{a}^T \hat{b})}} \left((J_{p_1} \hat{a})^T \hat{b} \right), \quad (6)$$

$$\nabla_{p_2} C_2(p_1, p_2, p_3) = -\frac{1}{\sqrt{1 - (\hat{a}^T \hat{b})}} \left((J_{p_2} \hat{a})^T \hat{b} + (J_{p_2} \hat{b})^T \hat{a} \right), \quad (7)$$

and

$$\nabla_{p_3} C_2(p_1, p_2, p_3) = -\frac{1}{\sqrt{1 - (\hat{a}^T \hat{b})}} \left((J_{p_3} \hat{b})^T \hat{a} \right). \quad (8)$$

Here, we only give the equation for $J_{p_1} \hat{a}$. The other jacobians are however similar.

$$J_{p_1} \hat{a} = \frac{I_3 - \left(\frac{p_1 - p_2}{|p_1 - p_2|}\right) \left(\frac{p_1 - p_2}{|p_1 - p_2|}\right)^T}{|p_1 - p_2|} \quad (9)$$

We make the choice of ϕ in Eq. (5) dependent on an overall smoothness parameter $k_s \in [0, 1]$. The smoothest possible boundary component ($k_s = 1$) is a regular N -polygon in which all internal angles have the same value $\phi = \frac{(N-2)}{N} \cdot \pi$. Hence, we set

$$\phi = \frac{(N-2)}{N} \cdot \pi \cdot k_s. \quad (10)$$

Entanglement of a boundary component is quantified through the *writhe* which can be viewed as the sum of all signed self-crossings averaged over all possible viewing directions. The writhe of a simple, closed, differentiable curve γ with points r_1 and r_2 along the curve is defined as the Gauss integral

$$Wr = \frac{1}{4\pi} \int_{\gamma} \int_{\gamma} dr_1 \times dr_2 \cdot \frac{r_1 - r_2}{|r_1 - r_2|^3}. \quad (11)$$

A similar definition could be used to calculate the writhe of a piecewise linear curve. However, it is computationally more efficient to consider an additional virtual curve γ' , to

calculate both the twist Tw and the Gauss linking number Lk of γ with γ' , and to employ the Călugăreanu-White-Fuller theorem [22]:

$$Wr = Lk - Tw \quad (12)$$

For details and a comparison of several methods to calculate the writhe, the reader is referred to [23]. The gradient of the writhe is computed using the formula from [24].

Now, we are able to define the writhe constraint function:

$$C_3(p) = |Wr(p)| - d_{Wr}, \quad (13)$$

where d_{Wr} denotes the maximum entanglement in an inequality constraint ($C_3(p) < 0$).

The writhe quantity has two drawbacks. For one thing, it is not defined for self-intersecting boundary components which we circumvent by imposing some additional distance constraints on non-adjacent points. For another, the writhe of any curve γ located on a sphere is zero. Therefore, boundary components tend to form non-planar configurations during uncoiling.

Alternatively, coiling out of the plane can be avoided by imposing a *planarity* constraint on the model. The cross product version of the so-called *shoelace formula* yields a vector n_p which is perpendicular to and whose length is twice the area of a given polygon. For non-planar polygonal chains such as *ABCMS*s, the result is an approximate "best-fit" normal vector:

$$n_p = \sum_{i=1}^N p_i \times p_{i+1} \quad (14)$$

with $p_{N+1} = p_1$. Let $\hat{n}_p = \frac{n_p}{|n_p|}$ be the unit normal, and let m_p be the mean of the boundary points. Then, a planarity constraint function can be defined as

$$C_4(p) = \frac{1}{N} \sum_{i=1}^N \underbrace{(\hat{n}_p^T (p_i - m_p))^2}_{d_i} - d_{Pl}, \quad (15)$$

limiting the mean squared distance of the points to a plane when used in an inequality constraint ($C_4(p) < 0$). The gradients with respect to the points are given by

$$\nabla_{p_j} C_4(p) = \frac{2}{N} \left[\sum_{i=1}^N d_i (J_{p_j} \hat{n}_p)^T (p_i - m_p) \right. \quad (16)$$

$$\left. - \sum_{i=1, i \neq j}^N \frac{1}{N} d_i \hat{n}_p + (1 - \frac{1}{N}) d_j \hat{n}_p \right] \quad (17)$$

with

$$J_{p_j} \hat{n}_p = \frac{1}{|n_p|} J_{p_j} n_p - \hat{n}_p \hat{n}_p^T J_{p_j} n_p \quad (18)$$

and

$$J_{p_j} n_p = \begin{pmatrix} 0 & p_{j+1z} - p_{j-1z} & p_{j-1y} - p_{j+1y} \\ p_{j-1z} - p_{j+1z} & 0 & p_{j+1x} - p_{j-1x} \\ p_{j+1y} - p_{j-1y} & p_{j-1x} - p_{j+1x} & 0 \end{pmatrix}. \quad (19)$$

Again, we make d_{Pl} in Eq. (15) dependent on an overall planarity parameter $k_{Pl} \in [0, 1]$. From some geometric considerations, we derive that the mean distance of the

boundary points to the fitted plane can not exceed $\frac{1}{16}$ of the boundary component's arc length L . Hence, we set

$$d_{Pl} \leq \left(\frac{L}{16} (1 - k_{Pl}) \right)^2. \quad (20)$$

The planarity constraint indeed helps to avoid coiling of the boundary component, but it sometimes also prevents uncoiling when the contour is already in an entangled state. Therefore, in practice, it makes sense to have both a writhe constraint and a planarity constraint, and to decrease the strength of the planarity constraint during uncoiling.

V. ACTIVE SKELETON MODELS

*ABCMS*s model the internal degrees of freedom of individual garment openings, but they do not incorporate the relationships between different boundary components. *Active Skeleton Models* extend the constraint-based concept to include a coarse representation of the overall geometry and topology of an article of clothing. We expect that the skeleton models could robustify tracking of the boundary components by providing some information about the relative poses of the openings. Furthermore, they can be used to define trajectories for the human limbs in dressing assistance.

We represent the skeletons as star-shaped structures with a single central point, one end point for each opening, and several points describing the paths from the central point to the boundary components. A suitable method for initializing a skeleton model together with the *ABCMS*s is described in Section VI. Possible deformations are again formalized by means of position-based constraints. We use internal angular and distance constraints to define the degrees of allowed stretching and bending. The attachments of the *ABCMS*s to the skeleton (Fig. 2) are characterized by the following two properties:

- 1) *Connectedness*: The main purpose of a skeleton model is to link the openings of a garment. Therefore, all skeleton end points are positioned close to the centers of their respective boundary components.
- 2) *Alignment*: The openings define the optimal entry directions into the garment interior. Hence, each skeleton end is roughly aligned with the normal vector of the boundary component it is attached to.

Connectedness is achieved by imposing a distance constraint on the skeleton end point q_1 and the boundary points p . Assuming that the mean m_p is a sufficient approximation of the boundary component center, we can define a simple constraint function

$$C_5(p, q_1) = |m_p - q_1| - d_c \quad (21)$$

with gradients

$$\nabla_{q_1} C_5(p, q_1) = - \frac{m_p - q_1}{|m_p - q_1|} \quad (22)$$

and

$$\nabla_{p_i} C_5(p, q_1) = \frac{m_p - q_1}{N |m_p - q_1|}, \quad (23)$$

where d_c is usually set to 0 in an equality constraint.

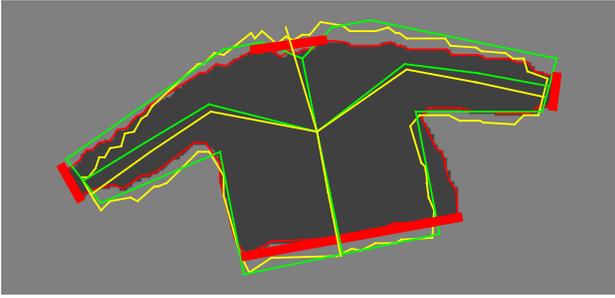


Fig. 3. Polygon-based model initialization. The segmented foreground region (dark gray) is approximated by a polygon (red), corrected for the camera perspective (yellow), and compared with a template (green). Minimization of a heuristic cost function yields the 2D projections of the garment openings (thick red lines). Additionally, a template skeleton (green) is deformed (yellow) to reflect the link structure between the openings.

Alignment is expressed through an angular constraint on the last skeleton segment $\overline{q_2q_1}$ and the unit normal \hat{n}_p of the boundary component:

$$C_6(p, q_1, q_2) = \arccos\left(\left(\hat{n}_p\right)^T \left(\frac{q_2 - q_1}{|q_2 - q_1|}\right)\right) - \phi_{Al} \quad (24)$$

The formulas for the gradients with respect to the boundary and skeleton points are similar to the gradients of the angular constraint function in Eq. (5), except that they include the jacobian of the unit normal from Eq. (18). The parameter ϕ_{Al} can be used in an inequality constraint ($C_6(p, q_1, q_2) < 0$) to set the misalignment tolerance.

VI. MODEL INITIALIZATION

In principle, the graph-based technique from [18] can be employed to initialize *ABCMS*. However, this method requires the openings to be fully visible to the sensor which makes it primarily applicable in the context of interactive perception. In the following, we describe an alternative initialization approach that exploits geometric and topological prior knowledge about clothing. Several sophisticated methods for garment unfolding [11]–[13] and flattening [14] exist. Therefore, it is reasonable to start from an item of clothing lying spread out on a tabletop. This assumption allows us to encode prior knowledge about different garment categories as 2D polygonal templates with points t_1, \dots, t_n , similar to [7] and [8]. In our framework, prior knowledge additionally includes the information about which polygon segments $\overline{t_a t_b}$ reflect openings. Furthermore, star-shaped template skeletons connecting the centers of these segments are given. Our polygon-based initialization method (Fig. 3) consists of the following five steps:

- 1) Hybrid foreground segmentation
- 2) Polygon matching
- 3) Heuristic search for openings
- 4) Skeleton deformation
- 5) Projection onto the tabletop

Hybrid foreground segmentation: We make use of both color and depth (point cloud) information for segmentation. As the pose of the tabletop plane is known or can be easily determined, we cut the calibrated point cloud in two parts

slightly above the tabletop and consider everything below as background and the points above as foreground. Due to noise and the fact that clothes are usually rather flat, this pre-segmentation is not accurate. Therefore, we only use it to learn *Gaussian mixture models* of the foreground and background colors, and to specify the seeds for the *GrabCut* segmentation algorithm [25] which is performed on a depth-registered color image.

Polygon matching: We approximate the foreground region by a polygon v_1, \dots, v_m using the method from [26]. The polygon is corrected for the camera perspective to make it commensurable with the template polygons. Then, a *turning function* based matching algorithm [27] is employed to compare the extracted polygon with each of the templates that were scaled to have equal arc lengths. The algorithm returns a metric reflecting the degree of shape similarity between the polygons which is used for garment classification.

Heuristic search for openings: As side products of the polygon matching algorithm, we obtain the rotation that was necessary to achieve the optimal match as well as a pair of corresponding reference points t_r and v_r . This makes it possible to lay the best-matching template on top of the extracted polygon in such a way that the overall orientations and the centroids match. Moreover, we are able to define the arc length $L(t_i v_j)$ as the absolute value of the sum of the (signed) arc lengths $L(t_i \widehat{t}_r)$ and $L(v_r \widehat{v}_j)$, possibly subtracted from the polygon length to obtain the minor arc length. Identifying the 2D projection of a garment opening means finding the vertices v_a and v_b of the extracted polygon that correspond to t_a and t_b in the best-matching template. We determine all approximately linear segments $\overline{v_i v_j}$ (i.e., the segments where the distances of intermediate vertices to the connecting line do not exceed a threshold). Then, we minimize a heuristic cost term F over all $\overline{v_i v_j}$ to find $\overline{v_a v_b}$:

$$F = \alpha_1 F_{arc} + \alpha_2 F_{dist} + \alpha_3 F_{orient} + \alpha_4 F_{length}, \quad (25)$$

where $F_{arc} = L(t_a \widehat{v}_i) + L(t_b \widehat{v}_j)$. F_{dist} corresponds to the Euclidean distance between the centers of $\overline{t_a t_b}$ and $\overline{v_i v_j}$. F_{orient} is the angle between $\overline{t_a t_b}$ and $\overline{v_i v_j}$, whereas F_{length} reflects their difference in length. The coefficients $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ control the relative influence of the subterms.

Skeleton deformation: Our method is capable of initializing a skeleton model along with the *ABCMS*. The algorithm deforms a copy of the given template skeleton so as to match the geometry of the extracted polygon. The central point remains unchanged because we assume it to be fixed with respect to the polygon centroid. Then, the individual skeleton branches are rotated and stretched in such a way that the end points match the centers of the segments $\overline{v_a v_b}$.

Projection onto the tabletop: Finally, the models are back-projected to the 3D world. In the case of the skeletons, this is a simple projection from the image plane onto the tabletop. The segments $\overline{v_a v_b}$ are converted to rectangular boundary components of height 1cm after projection. The violation of the smoothness assumption is going to vanish within the first iterations of constraint solving.

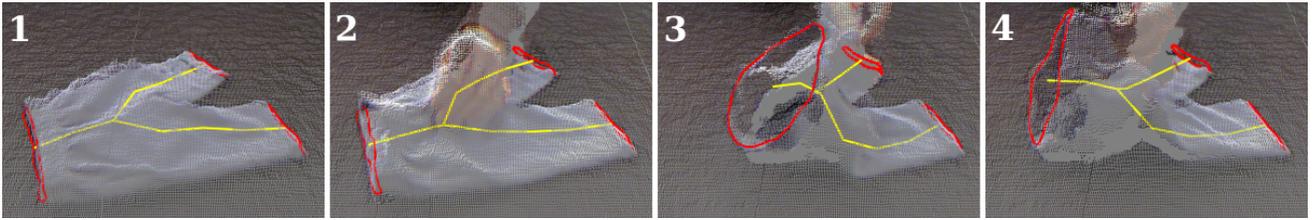


Fig. 4. Visual boundary component tracking of a pair of pants shortly after initialization (1), while grasped (2), lifted (3), and moved (4) by a human hand. The *ABCMs* are shown in red. The skeleton model is depicted in yellow.

VII. TRACKING *ABCMs* WITH POINT CLOUDS

As the constraint solving step ensures that the physical and topological assumptions are satisfied, the model points can be moved according to sensory input in a rather unrestricted way. In the present work, we focus on point cloud data from a *Kinect* sensor. Visually, the boundary components of garments appear as edges that, in the ideal case, form closed contours. 3D edge points of sufficiently thick clothes can be detected by employing the technique from Ückermann et al. [28] which finds differences in angle between adjacent surface normals in the point cloud. We apply the method twice, using slightly different parameters in the vicinity of the tabletop to detect the weak edges of flattened garments. At each frame, our algorithm collects a set E of relevant edge points by searching in a range around the *ABCM* points.

Subsequently, the model points are roughly aligned with the edge points in E . First, we determine a rigid transformation (translation and rotation) of the *ABCM* by using a reverse³ *Iterative Closest Point (ICP)* [29] approach. Then, the *ABCM* is deformed to match E even better. To this end, the 3D space is decomposed into *Voronoi* regions by the transformed model points before shifting the model points toward the centroids of the edge points that fall within their respective regions. This simple uninformed method neither considers any of our model assumptions nor does it update the skeleton estimates. Therefore, boundary component tracking (Fig. 4) relies on effective constraint solving. In our current implementation, occlusion is handled in an on-off way, i.e., visual tracking of an individual *ABCM* stops when another object crosses the view ray of a model point.

VIII. EXPERIMENTS

A. Initialization tests

In our experiments, we again considered three categories of clothing (leg warmers, pairs of pants, and sweaters). Our test set consisted of six different child garments, two of each category. We performed twelve initialization runs using the method described in Section VI. To this end, each item from the test set was spread out once on a black tabletop and once on a multicolored table cloth.

The algorithm classified the garment category correctly in all trials. Furthermore, we tested the initialization accuracy by comparing the results with the ground truth given by

³The ICP algorithm finds a transformation T from the points in E to the model points, and T^{-1} is applied to the model.



Fig. 5. Two trials from the initialization experiment. Ground truth segments provided by a human participant are shown in yellow. (a) A leg warmer in the black background condition. (b) A sweater in the multicolored background condition.

TABLE I
INITIALIZATION ERRORS WITH RESPECT TO GROUND TRUTH

	Black background		Multicolored background	
	Mean	SD	Mean	SD
Position error				
in pixels (image frame)	3.0	1.4	7.9	11.2
in mm (world frame)	7.4	2.9	18.9	23.2
Length error				
in pixels (image frame)	4.8	4.5	4.1	4.2
in mm (world frame)	12.2	10.5	10.6	10.5
Orientation error				
in degrees (image frame)	3.2	2.4	5.6	5.3
in degrees (world frame)	3.4	2.4	6.1	5.3

a human subject. The participant was presented with the same twelve depth-registered color images, and was asked to mark the openings of the garments (see Fig. 5). Afterwards, we measured the differences in position, length, and orientation between the algorithm-generated segments and ground truth (see Table I). In the black background condition, errors in position (7.4 mm), length (12.2 mm), and orientation (3.4 degrees) were small. While length errors did not differ much between the conditions, both position and orientation accuracy (mean error) and repeatability (standard deviation) suffered from the poorer segmentation results in the multicolored background condition.

B. Tracking evaluation

Point cloud based tracking was implemented using the *Image Component Library (ICL)*⁴ and runs in real time (30 Hz) on a PC with a modern graphics card. We used a knit cap, i.e., a garment with a single boundary component, to qualitatively

⁴<http://www.iclv.org/>

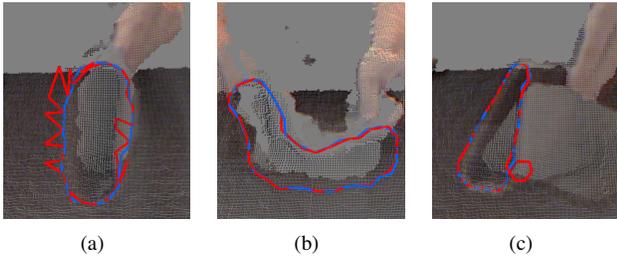


Fig. 6. Comparison of the ABCM tracking performance between several model configurations. (a) Unconstrained (red) vs. constrained (blue). (b) Without (red) vs. with entanglement constraints (blue). No stretching allowed. (c) Without (red) vs. with entanglement constraints (blue). Stretching allowed.

evaluate the influence of individual constraints on the tracking performance. In each trial, we initialized two *ABCMs* with different parameters using the method from [18], and observed how the models behaved during manipulation by a human. First, we tested an unconstrained model against an *ABCM* without stretching tolerance, $k_s = 0.5$, $k_{PI} = 0.85$, and a maximum writhes of 0.25. It can be clearly seen in Fig. 6(a) that the unconstrained model heavily violates our assumptions, while the constrained model reflects the real boundary component fairly well. In a second trial (6(b)), we compared the constrained *ABCM* from the first trial with an *ABCM* from which we removed the planarity and writhes constraints. We found that the performance of both models was similar and coiling occurred rarely, even in the condition without entanglement constraints. However, when we allowed stretching (up to 50 percent) in both conditions (Fig. 6(c)), we frequently observed boundary component entanglement if it was not explicitly constrained.

We also tested the dynamic behavior of *Active Skeleton Models* with attached *ABCMs* using the articles of clothing from our test set. We found that the models reacted plausibly to several manipulations such as grasping, lifting, moving, or slightly deforming parts of the garments. The models were indeed not robust against strong occlusions or deformations such as folding a sleeve, but we emphasize that there was no visual tracking of the skeletons or the overall garments. Taking into account that the skeleton models only followed the boundary component dynamics in a constraint-based manner, they represented the object poses surprisingly well.

C. Experiments with a bimanual robot

We investigated the performance of our model in a controlled robotic scenario that contained several elements of the dressing assistance task. A dual *Mitsubishi PA-10* arm with attached *Shadow Dexterous Hands* was available for the tests. The task of getting dressed consists of two alternating basic action patterns: pulling a garment part over a human limb, and pushing the limb through the garment interior. In our experiments, we focused on the latter, keeping in mind that the relative trajectory between the limb and the garment (see Section II) is the same in both cases. Specifically, the robot’s task was to slide a rod through a pant leg which can be regarded as an abstraction of dressing a leg prosthesis.

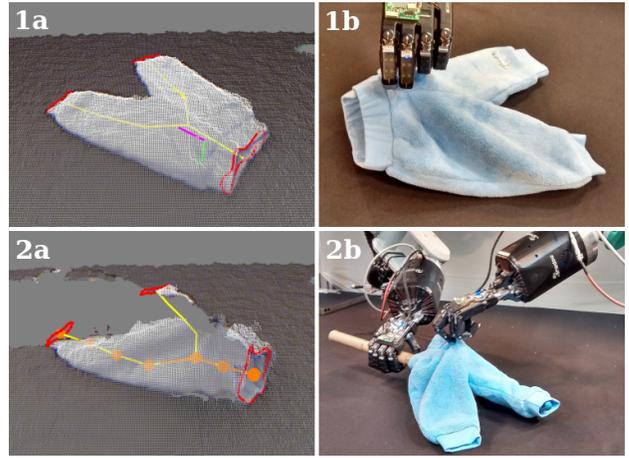


Fig. 7. Bimanual robot sliding a rod through a pant leg. (1a) Heuristic grasp pose detection. The approach vector is depicted in green, the orientation is shown in pink. (1b) The left hand lifting the garment in order to increase the size of the opening. (2a) Trajectory detection (orange) based on the skeleton model of the garment. (2b) The right hand pushing the rod through the garment interior. The tip of the rod follows the detected trajectory.

The task consisted of two steps (Fig. 7):

- 1) *Increasing the opening size*: We specified a heuristic grasp position (the highest point in a region behind the opening) which allowed the robot to slightly lift the garment from the tabletop in order to increase the size of the area circumscribed by the boundary component b_0 . Success was measured using the *relative opening size* $S \in [0, 1]$ which we defined as the ratio of the area to its upper bound (i.e., the area $\frac{L^2}{4\pi}$ of a circle):

$$S = \frac{2\pi|n_p|}{L^2}, \quad (26)$$

where L is the arc length of the boundary component and n_p is the normal vector from Eq. (14).

- 2) *Following the trajectory*: If the waist opening b_0 was wide enough ($S > 0.5$), the robot used the other hand to slide the rod through the boundary component and the interior of the garment toward the leg opening b_T . The rod was considered as an extension of the robot’s kinematic chain, and its tip followed a static path along the skeleton model while the orientation of the rod was aligned with the last segment of the trajectory.

The robot performed the task successfully with both pants from our test set. However, in the second run, the initial grasp was not strong enough, and the garment slipped out of the robot’s hand. Our system correctly detected the failure by checking the *relative opening size* S , triggered a new grasp, and completed the task. The time curves of S for the failed grasp and both successful grasps are shown in Fig. 8. Despite the lack of force control, the tip of the rod did not get stuck in the fabric. This indicates that the skeleton models provided suitable paths through the garment interior.

IX. CONCLUSION

The problem of robotic dressing assistance is quite under-researched. We speculate that the reason for this is the high

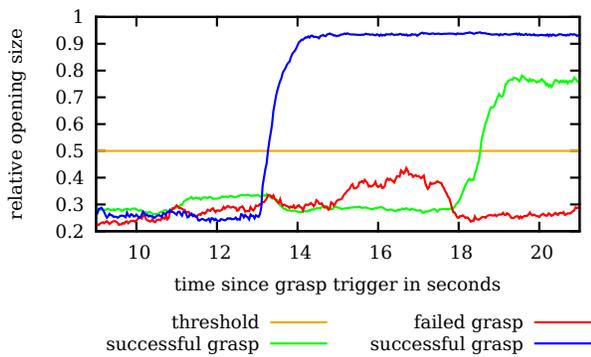


Fig. 8. Time curves of the *relative opening size* S of the boundary component b_0 (waist opening) during the robotic experiments.

complexity of the task which stems from the large number of subtasks that have to be solved. These subtasks include but are not limited to visual perception in a dynamic environment, modeling in a high-dimensional configuration space, integrating prior knowledge, and generating adequate action patterns. Rather than directly focusing on one of the sub-problems, we began the present paper with a consideration of the big picture. Our analysis revealed that the topological and functional properties of garments are highly correlated. In particular, the key role of the boundary components and their relationships became apparent. Therefore, we presented a novel constraint-based approach to modeling the dynamics of boundary components which we refer to as *ABCMs*.

We were able to show that *ABCMs* limit the degrees of freedom to a tractable level, and that they can be tracked visually using point cloud data. Furthermore, we suggested an initialization scheme incorporating prior knowledge about the article of clothing to be modeled. In experiments with a bimanual robot, we demonstrated the applicability of the proposed representation to the task of sliding a rod through a pant leg. In the general case, dressing assistance obviously demands a number of additional skills, both in perception and manipulation. These skills presumably require the integration of different modalities such as proprioception, touch, or force/torque sensing. We are optimistic that the independence of our model assumptions from any particular type of sensory input may prove beneficial in this regard.

ACKNOWLEDGMENTS

This work was partially funded by the German Research Council (DFG), grant EXC 277. We would like to thank Guillaume Walck, Robert Haschke, and everyone from the AGNI Grasp Lab for their helpful input and support.

REFERENCES

- [1] B. Willimon, I. Walker, and S. Birchfield, "A new approach to clothing classification using mid-level layers," in *ICRA*, 2013, pp. 4271–4278.
- [2] Y. Li, Y. Wang, M. Case, S.-F. Chang, and P. K. Allen, "Real-time pose estimation of deformable objects using a volumetric approach," in *IROS*, 2014, pp. 1046–1052.
- [3] B. Willimon, I. Walker, and S. Birchfield, "3d non-rigid deformable surface estimation without feature correspondence," in *ICRA*, 2013, pp. 646–651.

- [4] J. Schulman, A. Lee, J. Ho, and P. Abbeel, "Tracking deformable objects with point clouds," in *ICRA*, 2013, pp. 1130–1137.
- [5] Y. Kita, F. Kanehiro, T. Ueshiba, and N. Kita, "Clothes handling based on recognition by strategic observation," in *Humanoids*, 2011, pp. 53–58.
- [6] A. Ramisa, G. Alenyà, F. Moreno-Noguer, and C. Torras, "Determining where to grasp cloth using depth information," in *Int. Conf. of the Catalan Association for Artificial Intelligence*, 2011, pp. 199–207.
- [7] S. Miller, M. Fritz, T. Darrell, and P. Abbeel, "Parametrized shape models for clothing," in *ICRA*, 2011, pp. 4861–4868.
- [8] J. Stria, D. Prusa, V. Hlavac, L. Wagner, V. Petrik, P. Krsek, and V. Smutny, "Garment perception and its folding using a dual-arm robot," in *IROS*, 2014, pp. 61–67.
- [9] Y. Li, Y. Yue, D. Xu, E. Grinspun, and P. K. Allen, "Folding deformable objects using predictive simulation and trajectory optimization," in *IROS*, 2015, pp. 6000–6006.
- [10] S. K. Christian Bersch, Benjamin Pitzer, "Bimanual robotic cloth manipulation for laundry folding," in *IROS*, 2011, pp. 1413–1419.
- [11] A. Doumanoglou, A. Kargakos, T.-K. Kim, and S. Malassiotis, "Autonomous active recognition and unfolding of clothes using random decision forests and probabilistic planning," in *ICRA*, 2014, pp. 987–993.
- [12] Y. Li, D. Xu, Y. Yue, Y. Wang, S.-F. Chang, E. Grinspun, and P. K. Allen, "Regrasping and unfolding of garments using predictive thin shell modeling," in *ICRA*, 2015, pp. 1382–1388.
- [13] B. Willimon, S. Birchfield, and I. Walker, "Model for unfolding laundry using interactive perception," in *IROS*, 2011, pp. 4871–4876.
- [14] L. Sun, G. Aragon-Camarasa, S. Rogers, and J. P. Siebert, "Accurate garment surface analysis using an active stereo robot head with application to dual-arm flattening," in *ICRA*, 2015, pp. 185–192.
- [15] A. Colomé, A. Planells, and C. Torras, "A friction-model-based framework for reinforcement learning of robotic tasks in non-rigid environments," in *ICRA*, 2015, pp. 5649–5654.
- [16] T. Tamei, T. Matsubara, A. Rai, and T. Shibata, "Reinforcement learning of clothing assistance with a dual-arm robot," in *Humanoids*, 2011, pp. 733–738.
- [17] N. Koganti, J. G. Ngeo, T. Tomoya, K. Ikeda, and T. Shibata, "Cloth dynamics modeling in latent spaces and its application to robotic clothing assistance," in *IROS*, 2015, pp. 3464–3469.
- [18] L. Twardon and H. Ritter, "Interaction skills for a coat-check robot: identifying and handling the boundary components of clothes," in *ICRA*, 2015, pp. 3682–3688.
- [19] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.
- [20] M. Isard and A. Blake, "Condensation - conditional density propagation for visual tracking," *International Journal of Computer Vision*, vol. 29, no. 1, pp. 5–28, 1998.
- [21] M. Müller, B. Heidelberger, M. Hennix, and J. Ratcliff, "Position based dynamics," *Journal of Visual Communication and Image Representation*, vol. 18, no. 2, pp. 109–118, 2007.
- [22] F. B. Fuller, "The writhing number of a space curve," *Proceedings of the National Academy of Sciences*, vol. 68, no. 4, pp. 815–819, 1971.
- [23] K. Klenin and J. Langowski, "Computation of writhe in modeling of supercoiled dna," *Biopolymers*, vol. 54, no. 5, pp. 307–317, 2000.
- [24] R. de Vries, "Evaluating changes of writhe in computer simulations of supercoiled dna," *The Journal of chemical physics*, vol. 122, no. 6, p. 064905, 2005.
- [25] A. B. Carsten Rother, Vladimir Kolmogorov, "Grabcut: Interactive foreground extraction using iterated graph cuts," *ACM transactions on graphics*, vol. 23, no. 3, pp. 309–314, 2004.
- [26] X. C. He and N. H. C. Yung, "Corner detector based on global and local curvature properties," *Optical Engineering*, vol. 47, no. 5, 2008.
- [27] E. Arkin, L. Chew, D. Huttenlocher, K. Kedem, and J. Mitchell, "An efficiently computable metric for comparing polygonal shapes," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 13, no. 3, pp. 209–216, 1991.
- [28] A. Ückermann, R. Haschke, and H. Ritter, "Realtime 3d segmentation for human-robot interaction," in *IROS*, 2013, pp. 2136–2143.
- [29] P. Besl and N. D. McKay, "A method for registration of 3-d shapes," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 14, no. 2, pp. 239–256, 1992.