UC Berkeley
UC Berkeley Previously Published Works

Title
Modeling and Perception of Deformable One-Dimensional Objects

Permalink
https://escholarship.org/uc/item/36f4n4pp

Authors
Javdani, Shervin
Tandon, Sameep
Tang, Jie
et al.

Modeling and Perception of Deformable One-Dimensional Objects

Shervin Javdani, Sameep Tandon, Jie Tang, James F. O'Brien, Pieter Abbeel

Abstract— Recent advances in the modeling of deformable one-dimensional objects (DOOs) such as surgical suture, rope, and hair show significant promise for improving the simulation, perception, and manipulation of such objects. An important application of these tasks lies in the area of medical robotics, where robotic surgical assistants have the potential to greatly reduce surgeon fatigue and human error by improving the accuracy, speed, and robustness of surgical tasks such as suturing. However, different types of DOOs exhibit a variety of bending and twisting behaviors that are highly dependent on material properties. This paper proposes an approach for fitting simulation models of DOOs to observed data. Our approach learns an energy function such that observed DOO configurations lie in local energy minima. Our experiments on a variety of DOOs show that models fitted to different types of DOOs using our approach enable accurate prediction of future configurations. Additionally, we explore the application of our learned model to the perception of DOOs.

I. Introduction

The physical modeling of non-rigid objects like thread, rope, hair, and plants in real world settings is a challenging task. These objects are highly deformable, exhibiting bending and twisting behaviors, and their cross section is typically negligible in size compared to their length. Accurate physical modeling of these deformable one-dimensional objects (DOOs) has potential applications in, for example, graphics, robotic manipulation, and perception.

Our work is motivated by applications to medical robotics, where robotic surgical assistants, such as Intuitive Surgical's da Vinci, are increasingly being accepted in hospitals for surgery. These devices are usually tele-operated in a master-slave mode, and enable surgeons to accurately perform delicate operations at much finer scales than would be feasible with direct manual interaction. However, many surgical tasks are time consuming and require a high level of concentration, which can lead to surgeon fatigue.

Robotic surgical assistants present the possibility of autonomous execution of sub-tasks in surgical procedures. The automation of such sub-tasks would improve patient health by enhancing accuracy, reducing operation time, and reducing tedium and thus medical errors. Suturing, in particular knot-tying, stands out as a time-consuming, frequently recurring, and repetitive task which involves the perception and manipulation of suture material. Understanding the dynamics of DOOs is crucial for achieving high accuracy in this task.

Past work has often focused on graphical realism and has not emphasized accurate reproduction of specific physical instantiations. This work addresses the problem of modeling

The authors are with the Department of Electrical Engineering and Computer Sciences, UC Berkeley, CA 94720, U.S.A. Contact Email: sjavdani@eecs.berkeley.edu.
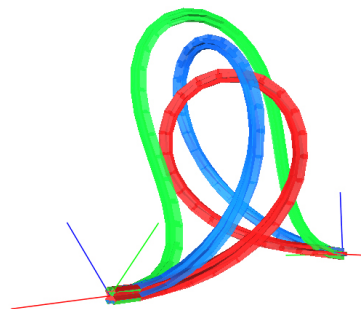

Fig. 1: Three simulated DOOs initialized with the same configuration exhibit different behavior under different model parameters.

specific physical instantiations of (DOOs), with high fidelity. By fitting model parameters to different types of DOOs from observed data, our approach accounts for differences in behavior due to underlying material properties (Figure 1).

The one-dimensional nature of DOOs makes the perception task challenging for two particular reasons. First, a DOO in an image is thin and easily occluded by objects in the foreground, including self occlusions. Second, many DOOs are uniform in color and texture, making it difficult to find correspondence points between images. We propose the application of a physically accurate model to address these challenges.

In this paper we present an approach for estimation of model parameters from real data. Our approach uses a generic energy function with free parameters within the energy model (which generalizes the one presented by Bergou et al. [1]). Using labeled training data consisting of different types of DOOs in many different 3D configurations, we learn the free parameters such that the configurations observed in the training data lie in local minima of the energy function. Our approach uses real training data to fit a simulation model specifically for different types of DOOs. Our experiments show that learning an energy function specific to the type of object yields improved prediction accuracy, and we demonstrate the effectiveness of our model and learned parameters for the task of perception.

Additional information can be found on our web site: `http://rll.berkeley.edu/ICRA_2011/`.

II. Related Work

We closely follow the work of Bergou et al. [1], [2], which uses an energy-minimization based approach for dynamic simulation of DOOs (which they refer to as elastic rods). This

approach accurately models DOO behavior, including twist-curvature interactions. Wakamatsu et al. [3], [4] also describe approaches for modeling of DOOs, and show experimental validation of model prediction based on observed behavior. Based on the demonstrated experimental results, we choose to build on the model of Bergou et. al, [1], as it provides a good combination of model accuracy and computational efficiency.

Kavraki and collaborators [5], [6], [7], [8], [9] consider the problem of motion planning for elastic objects, using a different energy formulation and model parameterization. Their model does not explicitly model twist-curvature interactions. During the planning step, they consider only quasistatic configurations, improving the robustness of the final trajectory. We adopt a similar quasistatic approach when learning models and evaluating performance. Their work primarily focused on challenges in motion planning and did not consider how to build object-specific dynamics models.

Sahalsto et al. [10] present a motion planning approach leveraging topological planning ideas, enabling them to tie knots—assuming the rope can be fixed in place at any point along its path. In particular, they make use of needle fixtures in order to create and maintain the integrity of loops throughout their knot tying procedure. Allowing fixtures during this step greatly simplifies the challenges in predicting how the rope will behave and enabled them to focus on addressing challenges in motion planning while side-stepping the challenges in perception and modeling.

Our work shares a common motivation with work in medical robotics where past work has addressed surgical skill modeling [11], [12], [13]. Mayer et al. [14], [15] show a surgical robot learning knot-tying from a demonstration. We are motivated by the goal of creating robotic surgical assistants which are faster, more robust, and more accurate than a human surgeon when executing surgical operations. In our past work [16], we demonstrated reliable execution of knot-tying maneuvers at superhuman speeds (using a learning from demonstration approach). However, in this work we did not attempt to model the dynamics of the suture, and accurate execution relied on having a consistent initial configuration of the suture and pickup point of the robot. The current work addresses the issue of accurately modeling specific types of suture and leveraging this for perception.

III. MODELING APPROACH

A. Preliminaries

Following Bergou et al. [1], we represent a DOO as a 3-tuple $\Gamma = (\mathcal{V}, \mathcal{E}, \mathcal{M})$, where $\mathcal{V} = \{v^0, \ldots, v^N\}$ is a collection of N 3-dimensional vectors representing vertices, $\mathcal{E} = \{e^0, \ldots, e^{N-1}\}$ is a collection of $N-1$ 3-dimensional vectors representing edges between the vertices $\mathcal{V}$, and $\mathcal{M} = \{(t^0, m_1^0, m_2^0), \ldots, (t^{N-1}, m_1^{N-1}, m_2^{N-1})\}$ is a collection of $N-1$ 3×3 matrices representing the material frames at each edge. We adopt the convention that quantities attached to vertices will be indexed by i, while quantities attached to edges will be indexed by j, for example vertex x^i and edge e^j. We define $t^j = \frac{e^j}{\|e^j\|}$ as the tangent vector pointing

in the direction of edge e^j. Together with m_1^j, m_2^j, these three vectors form an orthonormal basis denoting the material frame at edge e^j.

Let $(\kappa b)_i$ denote the curvature binormal at a vertex i:

$$(\kappa b)_i = \frac{2e^{j-1} \times e^j}{|e^{j-1}||e^j| + e^{j-1} \cdot e^j}$$

and let $\bar{l}_i = |e^{i-1}| + |e^i|$, so that $\bar{l}_i/2$ corresponds to the average length of the two segments closest to vertex x^i.

B. Constraints

In our model of DOOs, we enforce two separate types of constraints. The first, which we term inextensibility constraints, specify that each edge e^j is of fixed, uniform length. These constraints take the form:

$$e^j \cdot e^j - \bar{e}^j \cdot \bar{e}^j = 0 \quad \forall j = 1, \ldots, N-1$$

where $\bar{e}^j$ is the rest length of the jth edge.

In addition, we also enforce boundary point and frame constraints which fix the position of the start and end points as well as the material frame on the first and last edges:

$$(t^j, m_1^j, m_2^j) - (\bar{t}^j, \overline{m}_1^j, \overline{m}_2^j) = \mathbf{0}$$
$$v^i - \bar{v}^i = 0$$

for $i = \{0, N\}$ and $j = \{0, N-1\}$, where $\bar{v}^i$ is the constrained point and $(\bar{t}^j, \overline{m}_1^j, \overline{m}_2^j)$ is the constrained end frame. For a given DOO, Γ, we define the constraint matrix $C(\Gamma) = 0$ as the concatenation of the inextensibility and boundary constraints described above.

C. Energy Model

We adopt the energy model of Bergou et al. [1]. It models the energy of a specific configuration of Γ as a combination of bend energy and twist energy. Including a gravitational energy term results in:

$$E(\Gamma) = E_{\text{bend}}(\Gamma) + E_{\text{twist}}(\Gamma) + E_{\text{gravity}}(\Gamma).$$

The bending energy is defined as

$$E_{\text{bend}}(\Gamma) = \sum_{i=1}^{N} \frac{1}{2\bar{l}_i} \sum_{j=i-1}^{i} \omega_i^{jT} B \omega_i^j,$$

where $\omega_i^j = \left((\kappa b)_i \cdot m_2^j, -(\kappa b)_i \cdot m_1^j \right)^T$ for $j = \{i-1, i\}$ and B represents a 2×2 bending energy matrix.

To formulate a description of the twisting energy, we must first introduce the concepts of parallel transport, and the Bishop frame associated with each edge, which will allow us to define a zero-twist reference. The Bishop frames are defined to have their first axis aligned with e^j. To define the other two axes, we start by fixing the first Bishop frame (the one for edge e^0) by arbitrarily completing the axis aligned with e^0 into a right-handed frame. Then the second Bishop frame (the one for edge e^1) is defined as the frame that has its first axis aligned with e^1 and is minimally rotated with respect to the Bishop frame attached to e^0. The entire sequence of Bishop frames can be obtained this way and they define a zero-twist reference.

Formally, we let u^0 be a unit vector orthogonal to t^0. Then t^0, u^0, and $v^0 = t^0 \times u^0$ form a frame which we define as the Bishop frame at edge e^0. The remaining Bishop frames are defined by the relation $u^j = P_j(u^{j-1})$, where P_j is a rotation such that $P_j(t^{j-1}) = t^j$, $P_j(t^{j-1} \times t^j) = t^{j-1} \times t^j$.

The twist θ^j of a material frame $\mathcal{M}_j = (t^j, m_1^j, m_2^j)$ is defined as the rotation around t^j that aligns it with the Bishop frame (t^j, u^j, v^j). Hence the twist θ^j satisfies:

$$
\begin{array}{rcl}
m_1^j & = & \cos\theta^j \cdot u^j + \sin\theta^j \cdot v^j, \\
m_2^j & = & -\sin\theta^j \cdot u^j + \cos\theta^j \cdot v^j.
\end{array}
$$

The twist energy is defined in terms of the twist angles θ^j:

$$
E_{\text{twist}}(\Gamma) = \beta \sum_{j=1}^n \frac{(\theta^j - \theta^{j-1})^2}{\bar{l}_j}.
$$

Finally, the gravitational potential energy is given by

$$
E_{\text{gravity}}(\Gamma) = \mu \sum_{i=0}^n z^i \bar{l}_i
$$

where z^i represents the z-component of vertex v^i.

In the isotropic case ($B_j = \alpha I_{2 \times 2}$), the energy of a DOO reduces to:

$$
E(\Gamma) = \frac{1}{2} \sum_{i=1}^n \alpha \frac{(\kappa b)_i^2}{\bar{l}_i} + \frac{\beta(\theta^n - \theta^0)^2}{\sum_{i=1}^N \bar{l}_i} + \mu \sum_{i=0}^n z^i \bar{l}_i
$$

We let $\Theta = (B, \beta, \mu)$ be the vector of free parameters used to weight the different energy terms. In the following sections, we will often denote $E(\Gamma)$ as $E_\Theta(\Gamma)$ to make explicit the dependence of the energy function on Θ.

D. Finding Minimum Energy Configurations

Given a set of constraints C and an initial DOO configuration Γ, a common problem is to solve for the "closest" configuration[1] Γ' that locally minimizes the energy function E_Θ while satisfying C. Our rationale is that DOOs tend to move toward lower energy configurations. One application is predicting motion: given a configuration Γ_1 which we assume to be a local minimum of E_Θ under constraints C_1, and given a motion of the end points resulting in new constraints C_2, we could predict the new configuration of the DOO as a local minimum Γ_2 of E_Θ under C_2 which is found by starting a local minimization procedure from Γ_1. Another application is recovery of shape from noisy measurements of a DOO.

We use the following gradient-descent type scheme on the energy function $E_\Theta(\Gamma)$. We first take a step in the direction of the gradient of the energy with respect to the vertices $\mathcal{V}$ while holding twist angles fixed. Next, as the Bishop frames depend on the positions of the vertices, we compute the new Bishop frames. This gives us the reference frames for twist. Next we minimize the energy with respect to the twist angles. See [1] for a derivation of the analytic gradients.

To enforce constraints, we use the fast manifold projection method of [17], which alternates between the above energy minimization gradient step and a reprojection step using an augmented Lagrangian method. We iterate this procedure

until the improvement in the energy function from one step to another falls below some threshold.[2]

Let $\widetilde{\Gamma}_{\Theta,C}$ be the result of running the above energy minimization procedure starting from a configuration Γ using the energy function $E_\Theta(.)$ and the constraints C. We define the operator F as

$$
\widetilde{\Gamma}_{\Theta,C} = F(\Gamma, E_\Theta, C).
$$

E. Parameter Learning

The energy model described above has a number of free parameters Θ, namely the bending parameters matrix B (in the isotropic case, this is a scalar α), the twist angle parameter β, and the gravity/weight parameter μ. Because different types of DOOs possess different material properties, we expect that different combinations of parameters are better at predicting the behavior of different types of thread.

We now outline our two approaches to learning these parameters from training data. Both are based on the idea that under the right set of parameters Θ, observed training examples should lie in local minima of the corresponding energy function E_Θ.

Our first approach searches for a set of parameters Θ which minimizes the difference between a noisy observation of Γ and the estimated nearest-by local minimum of the energy. Concretely, let $\{(\Gamma^1, C^1), \ldots, (\Gamma^D, C^D)\}$ be a set of noisy training examples consisting of DOO configurations Γ^d together with associated end-point constraints C^d. In a slight abuse of notation, let $\mathcal{V}(\Gamma)$ refer to the vertices of the DOO configuration Γ, and let $\mathcal{M}(\Gamma)$ refer to the material frames. To quantify the difference between two configurations, we introduce the following metric:

$$
d(\Gamma^1, \Gamma^2) = \sum_{i=0}^N \frac{\|v^{i1} - v^{i2}\|_2}{N} + w_m \sum_{i=0}^{N-1} \frac{\|m_1^{i1} - m_1^{i2}\|_2}{N-1}
$$
$$
\text{where } \{(t^{id}, m_1^{id}, m_2^{id}) \mid i = 0, \ldots, N-1\} = \mathcal{M}(\Gamma^d)
$$
$$
\{v^{id} \mid i = 0, \ldots, N\} = \mathcal{V}(\Gamma^d). \tag{1}
$$

This metric is a weighted combination of two terms: a position error term, which is the l_2-norm of the difference between the vertices of the two configurations, and a material frame error, which is the l_2-norm of the difference between one axis of the material frames of the two configurations.

We propose to minimize the distance between the training data and our best estimate of the true configuration of the training data after energy minimization.

$$
\min_\Theta \ \sum_{i=1}^D d(\widetilde{\Gamma}_{\Theta,C^i}^i, \Gamma^i) \tag{2}
$$
$$
\text{s.t. } \widetilde{\Gamma}_{\Theta,C^i}^i = F(\Gamma^i, E_\Theta, C^i), \quad \forall i = 1, \ldots, D.
$$

---

[1]For example under the l_2 distance metric.

[2]This gradient descent method is a natural choice for quasi-static simulation. In the perception setting, however, we do not know the initial configuration, and start from a noisy estimate based on the images we want to improve upon. In this setting one might in principle instead be interested in the constrained local energy minimum closest to the initial configuration—rather than the one reached by descending the energy surface. However, it is unclear whether any scheme other than exhaustive search could find the exact solution and so we also use the described scheme for perception.

We call Equation (2) the static parameter learning objective.

Our second approach searches for a set of parameters Θ by considering simulation accuracy. Concretely, for each term in the objective we consider two training examples Γ^x and Γ^y which have been collected starting from Γ^x and ending in Γ^y. We then run a simulation in which we initialize with Γ^x and (potentially over several smaller steps) change the boundary constraints of Γ^x to the boundary constraints of Γ^y. We compare the predicted configuration with the actual configuration of Γ^y. This gives the following objective:

$$\min_\Theta \ \sum_{i=1}^{D} \sum_{j=1}^{D} w_{ij} d(\widetilde{\Gamma}^i_{\Theta,C^j}, \Gamma^j) \qquad (3)$$
$$\text{s.t. } \widetilde{\Gamma}^i_{\Theta,C^j} = F(\Gamma^i, E_\Theta, C^j) \ \ \forall i,j = 1,\dots,D.$$

The main difference between the two objectives can be seen in the term $\widetilde{\Gamma}^i_{\Theta,C^j}$, which is the result of applying the constraints from one configuration C^j (corresponding to configuration Γ^j) to another configuration Γ^i. The weights w_{ij} are specified beforehand, and allow us to specify the sets of motions which contribute to the objective function. We call Equation (3) the dynamic parameter learning objective.

These objective functions are nonlinear and non-convex. The dimensionality, however, is relatively low and in our experiments it was feasible to solve these using exhaustive search.

IV. APPLICATION TO PERCEPTION

Perception of DOOs can be particularly challenging due to a lack of salient features for stereo correspondence. A set of images is often consistent with multiple DOO configurations. An accurate DOO model can help disambiguate between these configurations, and improve detection reliability. For the remainder of this section, we focus on perception of isotropic materials.

A. Estimating Configuration

Given a set of camera images $\{I^1, I^2 \dots I^G\}$ of a DOO, our objective is to recover the most likely DOO configuration Γ. Such a configuration should be physically plausible while its reprojection onto the images agrees with visual evidence. We assess this likelihood as a weighted combination of a visual reprojection error and a DOO energy term. Concretely, assume each camera image I^c is represented as a binary map $I^c(x,y)$, where $I^c(x,y) = 1$ when Γ is detected on image I^c at pixel (x,y) (a detected pixel), and 0 otherwise. In practice, we use a Canny edge detector with a blurred, offset Gaussian kernel to generate binary images I^c. We solve for the optimal Γ by minimizing the following objective function:

$$E_{\text{vis}}(\Gamma) = E(\Gamma) + \nu \sum_{i=1}^{G} E^i_{\text{proj}}(\Gamma).$$

Here ν is a weighting constant, and $E^i_{\text{proj}}(\Gamma)$ is the visual reprojection error of configuration Γ on camera image i. We compute visual reprojection error by projecting every edge $e^j \in \mathcal{E}(\Gamma)$ onto each camera image and summing the distances to the nearest detected pixels.

Our full perception algorithm incrementally constructs possible DOO configurations edge by edge (Algorithm 1).

Algorithm 1 Perception for DOOs

Initialize $\mathcal{S}_1 = \{\Gamma_1\}$, C
for all $i = 1 \dots N$ **do**
  for all $\widehat{\Gamma}_i \in \mathcal{S}_i$ **do**
    Generate set of new hypotheses $\{\widehat{\Gamma}_{i+1}\}$ from $\widehat{\Gamma}_i$
    Solve $\{\widetilde{\Gamma}_{i+1}\} = F(\{\widehat{\Gamma}_{i+1}\}, E_{\text{vis}}, C)$
    Add $\{\widetilde{\Gamma}_{i+1}\}$ to $\mathcal{S}_{i+1}$
  end for
  Select H hypotheses from $\Gamma_{i+1} \in \mathcal{S}_{i+1}$
end for
return $\arg \min_{\Gamma_N} E_{\text{vis}}(\Gamma_N)$, $\Gamma_N \in \mathcal{S}_N$

To recover a DOO with a fixed number of vertices N, we pick an initial hypothesis Γ_1 with vertex v^0 and edge e^0, and constraint set C which includes the inextensibility and start constraints. At each step, we define $\mathcal{S}_i$ to be the set of configuration hypotheses with i edges. For each configuration in $\mathcal{S}_i$, we generate several new hypotheses by adding an edge e^{i+1} using visual reprojection error alone. Specifically, we densely sample edges e^{i+1} such that $e^i \cdot e^{i+1} > 0$. Next, we sort the edges based on the visual reprojection error. We then select the edge e^j with the lowest error, and remove all edges within a certain angle of e^j. We repeat this process until all edges have been selected or removed. For each new hypothesis, we solve for a more likely configuration by minimizing $E_{\text{vis}}(\Gamma_{i+1})$.

After minimizing, we prune hypotheses as follows: we sort all hypotheses based on the score E_{vis}. Next, we greedily keep the hypothesis with the lowest score, and remove all hypotheses which are too similar based on a function of the distance between the corresponding last vertex v^{i+1} and the angle between the corresponding last edge e^i. We repeat until we have selected up to H hypotheses. After N edges, we return the hypothesis with the lowest score $E_{\text{vis}}(\Gamma_N)$.

B. Estimating Twist Angles

As our algorithm is building up the DOO incrementally, it is not including an end-point or end-frame constraint in the energy minimization. As a byproduct, the energy minimum will find zero twist energy for isotropic materials. To recover an estimate of the twist angles, after having run the perception algorithm, we propose to find the end-frame constraint, and corresponding twist angles, for which the recovered location of the vertices is closest to a local minimum of the energy E.

Concretely, we are given D configurations $\Gamma^1,\dots,\Gamma^D$ with associated constraints[3] $C^1,\dots,C^D$. We let ϕ be the free parameter we try to recover, corresponding to not knowing how much the end frame was rotated around the twist axis

---

[3]In practice, we initialize the end-frame constraint of C^i with the last bishop frame of Γ^i.

relative to the zero twist case.[4] A particular choice of ϕ induces a corresponding set of constraints $T_\phi(C)$ as follows: Given constraints C with end-frame constraint $(\bar{t}, \overline{m}_1, \overline{m}_2)$, let $T_\phi(C)$ be the same set of constraints but with the end frame rotated by ϕ about $\bar{t}$. We propose to recover ϕ by minimizing the following twist estimation objective:

$$\min_\phi \ \sum_{i=1}^{D} d(\widetilde{\Gamma}^i_{\Theta, T_\phi(C^i)}, \Gamma^i) \qquad (4)$$
$$\text{s.t. } \widetilde{\Gamma}^i_{\Theta, T_\phi(C^i)} = F(\Gamma^i, E_\Theta, T_\phi(C^i)), \quad \forall i = 1, \ldots, D.$$

V. EXPERIMENTS

A. Data Gathering Setup

To gather training data for our parameter learning and perception experiments, we used a calibrated 3-camera rig (Figure 2) to gather images of several types of DOOs: three different types of isotropic suture, namely Ethicon coated vicryl polyglactin 910 1 suture (vicryl), Oasis nylon MV-N666 monofilament 1 suture (nylon), and Ethicon silk K883 2-0 suture (silk), and one type of anisotropic gift wrapping ribbon, Hallmark 249RMS6006 (approximately 6.7mm wide). We prepared an example of each type of DOO of length 99mm, and hand-painted marks every 3mm (for a total of 34 vertices). We fixed the end points and frames of each DOO using alligator clips. For each DOO, we varied the distance between end points to be $30, 42, 54, 66, 78, 90$mm, for a total of six end-point distances. For each end-point distance, we fixed the start frame and varied the rotation of the end frame to be $0, 30, ..., 270°$ for a total of ten end-frame rotations. In total, we had three images for each of six end-point distances and 10 end-frame rotations for a total of 60 different configurations per DOO type. To extract ground truth DOO configurations, we hand-labeled each of the 34 marks in the 3 images for each configuration, and used triangulation to estimate the corresponding 3-D position. The material frames at each end point are known. For anisotropic DOOs we hand-labeled the material frame at each edge by labeling additional markers in the image.

B. Isotropic Material, Static Objective

We evaluate our parameter learning algorithm using the static objective of Equation (2) on our three different isotropic materials (sutures).[5] For each type of DOO, we first constructed a visualization of the static objective function over a densely sampled subset of parameter space, which we term the static error landscape (Figure 3).[6] These plots make a strong case for performing parameter learning. Finding

[4]For simplicity of exposition, we assume only one free parameter ϕ, which is the case when we know the end-constraint motions between $C^1, \ldots, C^D$, which would be the case if a robot were to manipulate the suture through $C^1, \ldots, C^D$ without re-grasping. If not the case, we would end up with an additional parameter ϕ for each re-grasp (assuming no knowledge about what happened to the twist between grasps) and the problem would break down into several disjoint problems with a single parameter ϕ.

[5]For isotropic materials we set $w_m = 0$, as our data does not include labeled material frames. Even without the material frame data for training, our approach successfully predicts DOO configurations to within 3.2mm.

[6]Since we can scale all our parameters by a constant without altering the minimum energy configurations, we fix μ while varying α and β
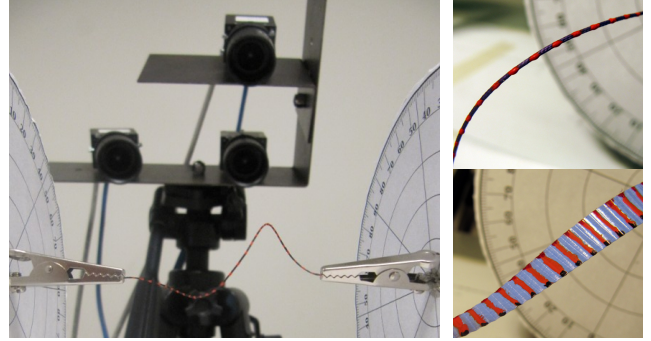

Fig. 2: Left: Data gathering setup. Top Right: Vicryl suture. Bottom Right: Ribbon material. We place orange marks across the DOOs for stereo correspondence.

	Nylon Test (mm)	Silk Test (mm)	Vicryl Test (mm)
Nylon	1.40 ± 0.09	3.09 ± 0.17	$\mathbf{2.14 \pm 0.16}$
Silk	1.59 ± 0.09	$\mathbf{2.56 \pm 0.19}$	2.73 ± 0.17
Vicryl	$\mathbf{1.36 \pm 0.09}$	3.201 ± 0.17	2.15 ± 0.16

TABLE I: The effect of varying DOO type on simulator accuracy. The rows correspond to the type of DOO used in the training set, while the columns show the type of DOO used in the test set. We use leave-one-out cross validation with Equation (2) as the test score and report the mean and standard error over all configurations in our data set. Bolded entries correspond to the best learned parameters for each test set.

optimal settings of the parameters can greatly improve the accuracy of the simulation.

To compare performance across different types of DOOs, we varied the type of DOO used in the training set together with the type of DOO used in the test set, making a total of 9 different test setups. For each test setup, we performed leave-one-out cross validation, training on 59 configurations and testing on the 60th. We performed the minimization over material parameters using an exhaustive search. Using Equation (2) as the test score, we report the mean and standard errors over all configurations in our data set in Table I. Our approach was able to distinguish the material properties of the silk suture from those of the nylon and vicryl—the learned parameters for the silk suture significantly outperform all other parameters on the silk test set. The learned parameters from both the nylon and vicryl training sets performed well on both nylon and vicryl test sets. The resulting errors for nylon and vicryl suture are always within 1 standard error, implying that their behavior is similar when aggregating over the entire data set.

When we break down the data further by the configuration's end-point distance in addition to the type of DOO, we observe that learning different parameters for different parts of the configuration space can improve performance. Figure 4 shows how combining data with different end-point distances averages out individual trends in the error landscapes specific to different parts of the configuration space. Table II shows the effect of varying DOO type and end-point distance

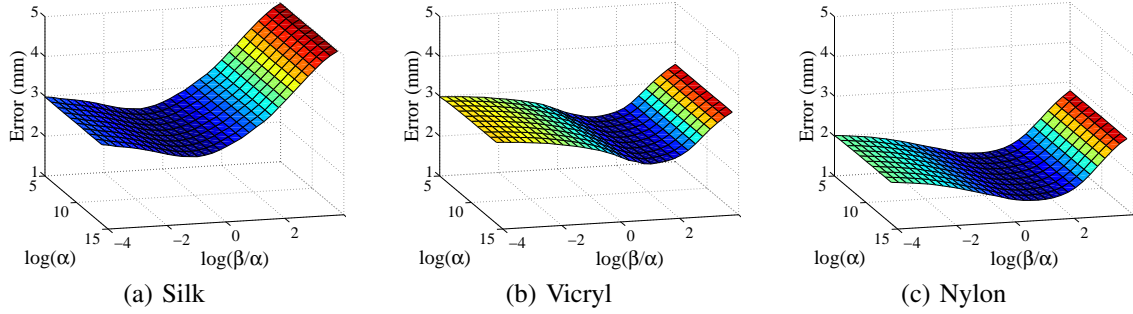(a) Silk   (b) Vicryl   (c) Nylon

Fig. 3: Static error landscapes for isotropic DOOs. We use Equation (2) as the error and show the error landscapes for silk, vicryl, and nylon DOOs. The parameters with lowest test error are different for each type of DOO.

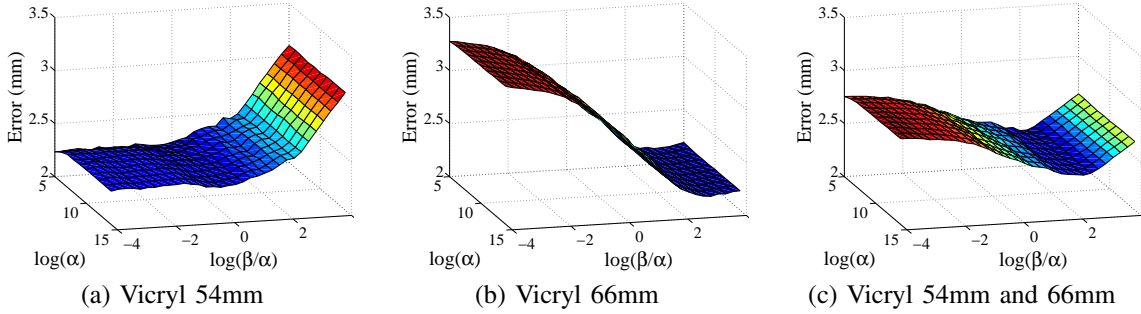(a) Vicryl 54mm   (b) Vicryl 66mm   (c) Vicryl 54mm and 66mm

Fig. 4: Example of aggregating data by end-point distance on simulator performance. We use Equation (2) as the error. The results for the 54mm data suggest that the optimal parameters lie toward lower $\frac{\beta}{\alpha}$ ratios, while the results for the 66mm data suggest the optimal parameters lie toward higher $\frac{\beta}{\alpha}$ ratios. Combining the data sets averages out the individual trends. While figures shown are for vicryl suture, other materials exhibit similar behavior.

on simulator accuracy. To construct our training sets, we selected 3 different end-point distances (30mm, 54mm, 78 mm) for all 3 types of DOOs. We created our test sets in the same manner. We use leave-one-out cross validation with Equation (2) as the test score. In 6 of the 9 test sets, the minimum test score is achieved when learning and testing on data at the same end-point distance (corresponding to the diagonal entries of Table II). Even when the minimum test score is not on the diagonal, the diagonal entry is always within 7% of the minimum. In general, partitioning by end-point distance and DOO type outperforms partitioning by DOO type alone. These results implies that while a global choice of parameters already provides accurate predictions, richer models have the potential to further improve simulation and learning parameters local to a region in the configuration space could be a plausible way to obtain such richer models.

C. Isotropic Material, Dynamic Objective

In this section we present results comparing the performance of our parameter learning algorithm using the static objective of Equation (2) versus the dynamic objective of Equation (3).[7] We partitioned our data set by the end-point distance as well as the type of DOO. The results are shown in Table III. Dynamic parameters (those learned on the dynamic objective) perform about as well as static parameters (those

learned on the static objective) on the static test set. However, dynamic parameters perform significantly better than static parameters on the dynamic test set. This result suggests that the dynamic objective is more informative than the static objective to estimate material parameters.

D. Anisotropic Material

We also explored the performance of our parameter learning algorithm under both static and dynamic objectives on anisotropic DOOs.[8] In both cases, we restrict our bending energy matrix B to be diagonal, yielding two parameters α_1, α_2. We fix both the gravity parameter μ and the twist parameter β, and varied the bend parameters α_1, α_2. Note that for the specific ribbon material we test on, it is very difficult to bend it about one axis. Therefore, we expect one bend coefficient to be significantly larger than the other.

Figure 5 (a) and (b) show static and dynamic error landscapes over a densely sampled subset of parameter space. Three distinct behavior regimes can be seen in this figure. Figure 5 (b) illustrates how our dynamic objective helps avoid local minima in parameter space which look reasonable under the static objective.

E. Estimating Configuration

We applied our perception algorithm to all images from our dataset, with parameters Θ from our static learning, $\nu =$

---

[7]Here we set w_{xy} to be nonzero iff Γ^x, Γ^y have the same end-frame distances and end-frame rotations which differ by $120°$.

[8]In Equation (1), we set $w_m = 6.7$, the width of our ribbon in mm.

	Nylon 30mm	Nylon 54mm	Nylon 78mm	Silk 30mm	Silk 54mm	Silk 78mm	Vicryl 30mm	Vicryl 54mm	Vicryl 78mm
Nylon 30mm	**1.02 ± 0.07**	1.13 ± 0.11	1.71 ± 0.27	2.61 ± 0.30	3.70 ± 0.42	3.27 ± 0.56	1.93 ± 0.27	2.61 ± 0.44	2.35 ± 0.50
Nylon 54mm	1.05 ± 0.08	**1.08 ± 0.10**	1.65 ± 0.28	2.32 ± 0.21	3.27 ± 0.32	3.25 ± 0.60	**1.81 ± 0.30**	2.34 ± 0.39	2.47 ± 0.50
Nylon 78mm	1.43 ± 0.17	1.25 ± 0.10	1.65 ± 0.29	1.99 ± 0.30	2.33 ± 0.20	**2.97 ± 0.56**	2.18 ± 0.27	2.16 ± 0.34	2.72 ± 0.39
Silk 30mm	1.78 ± 0.25	1.39 ± 0.09	1.59 ± 0.29	**1.85 ± 0.29**	1.88 ± 0.18	**2.97 ± 0.51**	2.34 ± 0.24	2.24 ± 0.33	2.93 ± 0.32
Silk 54mm	2.05 ± 0.26	1.56 ± 0.11	1.63 ± 0.29	**1.85 ± 0.30**	**1.74 ± 0.19**	3.04 ± 0.47	2.41 ± 0.23	2.23 ± 0.31	3.04 ± 0.29
Silk 78mm	1.61 ± 0.20	1.32 ± 0.09	**1.58 ± 0.29**	**1.85 ± 0.29**	2.14 ± 0.20	3.16 ± 0.49	2.25 ± 0.26	2.28 ± 0.34	2.91 ± 0.33
Vicryl 30mm	1.15 ± 0.09	**1.08 ± 0.10**	1.66 ± 0.30	2.56 ± 0.21	3.09 ± 0.24	3.11 ± 0.56	1.83 ± 0.26	2.06 ± 0.32	2.44 ± 0.50
Vicryl 54mm	2.17 ± 0.32	1.66 ± 0.18	1.66 ± 0.30	2.07 ± 0.30	2.17 ± 0.28	3.52 ± 0.35	1.95 ± 0.20	**1.85 ± 0.29**	3.04 ± 0.30
Vicryl 78mm	2.37 ± 0.52	2.31 ± 0.61	1.96 ± 0.28	5.25 ± 0.89	5.56 ± 0.89	3.45 ± 0.51	3.52 ± 0.59	3.87 ± 0.76	**2.20 ± 0.56**

TABLE II: The effect of varying DOO type and end-point distance on simulator accuracy. The rows correspond to the type and end-point distance of DOO used in the training set, while the columns show the type and end-point distance of DOO used in the test set. We use leave-one-out cross validation with Equation (2) as the test score and report the mean and standard error over all configurations in our data set. All results are reported in mm.

Vicryl	30mm Static	54mm Static	78mm Static	30mm Dynamic	54mm Dynamic	78mm Dynamic
30mm Static	**1.81 ± 0.42**	2.69 ± 0.27	3.33 ± 0.55	4.57 ± 0.40	6.92 ± 0.60	10.43 ± 0.39
54mm Static	1.98 ± 0.24	**2.36 ± 0.26**	3.40 ± 0.37	6.72 ± 0.53	8.84 ± 0.78	10.88 ± 0.24
78mm Static	3.99 ± 0.61	5.02 ± 0.62	**3.09 ± 0.68**	5.57 ± 0.56	10.73 ± 0.81	11.13 ± 0.94
30mm Dynamic	2.17 ± 0.41	3.01 ± 0.16	3.42 ± 0.42	**1.50 ± 0.19**	4.81 ± 0.34	8.59 ± 0.30
54mm Dynamic	1.99 ± 0.47	3.11 ± 0.25	3.38 ± 0.49	2.03 ± 0.37	**4.67 ± 0.47**	8.44 ± 0.32
78mm Dynamic	1.95 ± 0.48	3.47 ± 0.36	3.29 ± 0.56	3.09 ± 0.45	5.11 ± 0.50	**8.43 ± 0.37**

TABLE III: The effect of varying objective and end-point distance on simulator accuracy. Results are reported for vicryl suture. The rows correspond to the objective and end-point distance used in the training set, while the columns show the objective and end-point distance used in the test set. We use leave-one-out cross validation with Equation (2) and Equation (3) as the test score for static and dynamic objective respectively, and report the mean and standard error over all configurations in our data set. We restrict the test configurations of the static objective to those that are used in the dynamic objective. All results are reported in mm. The minimum test score is found along the diagonal in all 6 of 6 test sets (i.e. where the best parameters found on a given training set perform the best on the corresponding test set).

(a) Ribbon, static objective   (b) Ribbon, dynamic objective   (c) Ribbon Behavior
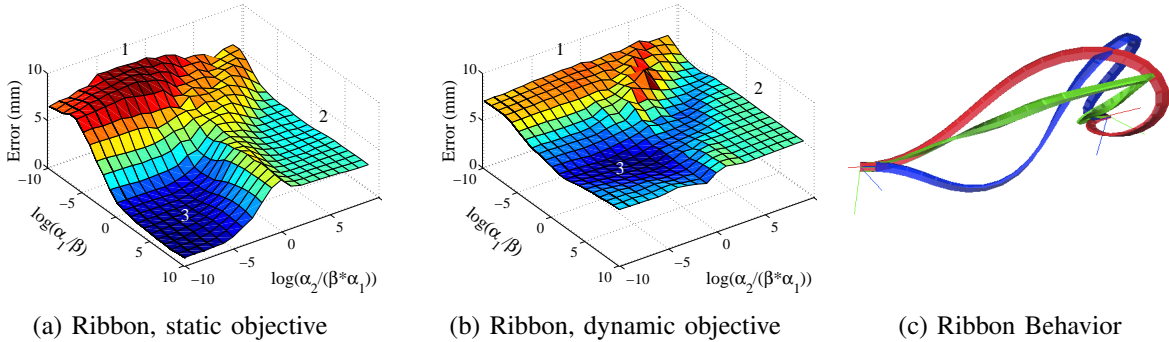
Fig. 5: Static and dynamic error landscapes for anisotropic DOOs. We use Equation (2) and Equation (3) as the error for (a) and (b) respectively. We fix β to 2.0 and vary bend parameters α_1 and α_2. Both plots show roughly three distinct behavior regimes: 1) $\beta > \alpha_1, \alpha_2$, 2) $\alpha_2 > \alpha_1 > \beta$, and 3) $\alpha_1 > \alpha_2 > \beta$. Compared to plot (a), plot (b) shows clear minima in regions 2) and 3). Figure (c) shows sample DOOs with identical end-point and end-frame constraints, but different parameters. Colors red, green, and blue corresponding to regions 1, 2, and 3 respectively. The difference in parameters results in a wide variation of the final configurations. For region 1, the ribbon makes sharp bends in order to avoid twisting. For region 2, the ribbon predicts bends about the "wrong" axis. For region 3, the ribbon uses twist to avoid bending about the wrong axis, which matches reality most accurately.

0.005, and $H = 40$. In order to distinguish between errors due to missing gaps from our Canny detection and the error minimization, we filled in gaps based on hand-labeled data. Human labels were used to seed the algorithm with an initial guess of start vertex v^0 and edge e^0.

When only considering the visual reprojection error (E_{proj}), multiple configurations seem plausible (Figure 6).

Using our approach, we were able to recover the correct suture configuration on all 180 test sets. To assess reconstruction error, we evaluate the average l_2-norm between detected vertices and hand-labeled data. Our approach was highly accurate, with errors of 0.023mm for nylon, 0.028mm for silk, and 0.040mm for vicryl.
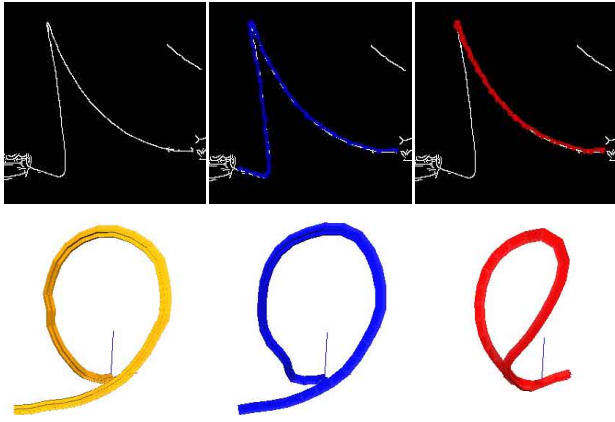
Fig. 6: Top: Canny images with reprojections. Bottom: 3D configurations. Yellow is ground truth, blue is the correct configuration, and red is another plausible configuration when considering the visual projection only. By considering model energy in addition to visual reprojection, we correctly choose the more physically plausible blue configuration.
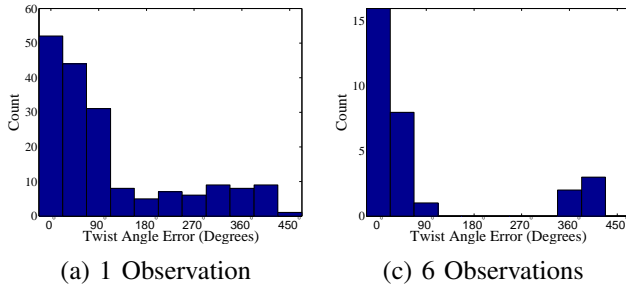

(a) 1 Observation   (c) 6 Observations

Fig. 7: Histograms of angle errors of end-frame rotation estimation across all thread types.

F. Estimating Twist Angles

We evaluated our twist estimation scheme from Section IV-B. Using configurations recovered with our perception algorithm, we evaluate the twist estimation objective in Equation (4) by searching for ϕ over the interval $-2\pi \leq \theta \leq 2\pi$ at $\frac{\pi}{12}$ increments. For our energy minimization operator $F(\Gamma, E_\Theta, C)$ we used the Θ learned under the static objective and partitioned by end-point distance.

Figure 7 shows the effect of varying the number of observations D. Our results suggest that certain configurations of vertices can be well-explained by a range of twist angles. Increasing the number of observed configurations increased estimation accuracy. With only one observation, we had a 53.3% success rate of predicting end-frame rotation to within $90°$. With 6 observations, this success rate was 80%.

VI. CONCLUSION AND FUTURE WORK

We have presented a method for learning adaptive models of deformable one-dimensional objects from data. Our experiments confirm that our method produces learned models in which minimum energy configurations better reflect the actual observed data. We applied these learned models to the perception of suture, demonstrating their applicability

for finding both 3D configurations and twist angles. Going forward, we plan to investigate using the learned models for motion planning and controlled manipulation of suture with visual feedback.

REFERENCES

[1] M. Bergou, M. Wardetzky, S. Robinson, B. Audoly, and E. Grinspun, "Discrete elastic rods," *ACM Trans. Graph.*, vol. 27, pp. 63:1–63:12, August 2008.

[2] M. Bergou, B. Audoly, E. Vouga, M. Wardetzky, and E. Grinspun, "Discrete viscous threads," *ACM Trans. Graph.*, vol. 29, pp. 116:1–116:10, July 2010.

[3] H. Wakamatsu and S. Hirai, "Static modeling of linear object deformation based on differential geometry," *International Journal of Robotic Research*, vol. 23, no. 3, pp. 293–311, 2004.

[4] H. Wakamatsu, K. Takahashi, and S. Hirai, "Dynamic modeling of linear object deformation based on differential geometry coordinates," in *Proceedings of the International Conference on Robotics and Automation (ICRA)*, 2005.

[5] F. Lamiraux and L. E. Kavraki, "Planning paths for elastic objects under manipulation constraints," *International Journal of Robotic Research*, vol. 20, no. 3, pp. 188–208, 2001.

[6] E. Anshelevich, S. Owens, F. Lamiraux, and L. Kavraki, "Deformable volumes in path planning applications," in *Proceedings of the International Conference on Robotics and Automation (ICRA)*, 2000.

[7] C. Holleman, L. Kavraki, and J. Warren, "Planning paths for a flexible surface patch," in *Proceedings of the International Conference on Robotics and Automation (ICRA)*, 1998.

[8] L. Kavraki, F. Lamiraux, and C. Holleman, "Towards planning for elastic objects," in *Robotics: The Algorithmic Perspective*, 1998.

[9] M. Moll and L. E. Kavraki, "Path planning for deformable linear objects," *IEEE Trans. on Robotics*, vol. 22, no. 4, pp. 625–636, Aug. 2006.

[10] M. Saha, P. Isto, and J.-C. Latombe, "Motion planning for robotic manipulation of deformable linear objects," in *Proceedings of the International Symposium on Experimental Robotics (ISER)*, 2006.

[11] T. Kowalewski, J. Rosen, L. Chang, M. Sinanan, and B. Hannaford, "Optimization of a vector quantization codebook for objective evaluation of surgical skill," *Studies in Health Technology and Informatics*, 2004.

[12] J. Rosen, B. Hannaford, C. Richards, and M. Sinanan, "Markov modeling of minimally invasive surgery based on tool/tissue interaction and force/torque signatures for evaluating surgical skills," *IEEE Trans. on Biomedical Engineering*, 2001.

[13] H. Lin, I. Shafran, D. Yuh, and G. Hager, "Towards automatic skill evaluation: detection and segmentation of robot-assisted surgical motions," *Computer Aided Surgery*, 2006.

[14] H. Mayer, I. Nagy, A. Knoll, E. Braun, and R. Bauernschmitt, "Adaptive control for human-robot skill transfer: trajectory planning based on fluid dynamics," in *Proceedings of the International Conference on Robotics and Automation (ICRA)*, 2007.

[15] H. Mayer, I. Nagy, D. Burschka, A. Knoll, E. Braun, R. Lange, and R. Bauernschmitt, "Automation of manual tasks for minimally invasive surgery," in *Proceedings of the International Conference on Autonomic and Autonomous Systems (ICAS)*, 2008.

[16] J. van den Berg, S. Miller, D. Duckworth, H. Hu, A. Wan, X.-Y. Fu, K. Goldberg, and P. Abbeel, "Superhuman performance of surgical tasks by robots using iterative learning from human-guided demonstrations," in *Proceedings of the International Conference on Robotics and Automation (ICRA)*, 2010.

[17] R. Goldenthal, D. Harmon, R. Fattal, M. Bercovier, and E. Grinspun, "Efficient simulation of inextensible cloth," in *Proceedings of ACM SIGGRAPH 2007*, 2007.