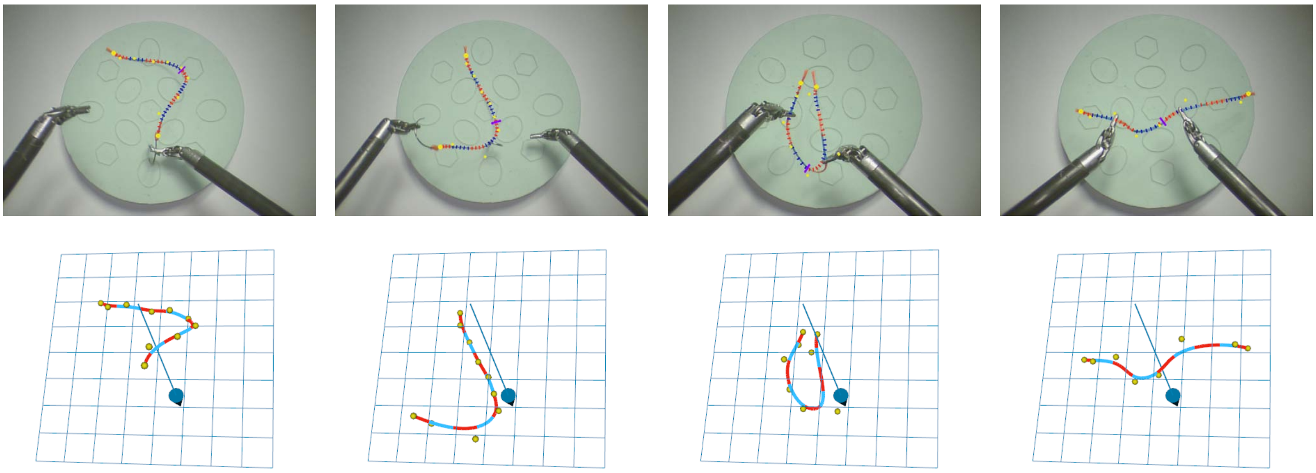
Deformable Tracking of Textured Curvilinear Objects

Nicolas Padoy
padoy@jhu.edu
Gregory Hager
hager@jhu.edu

Johns Hopkins University
Baltimore, Maryland, USA



Thread tracking illustration on two sequences: original images overlaid with spline models (top); virtual views (bottom).

Threads and wires are deformable 3-dimensional (3D) and curvilinear objects which are commonly manipulated by humans in various medical and manufacturing tasks. Several applications, including computer-assisted evaluation, augmented reality guidance, and autonomous robotic manipulation [2, 3] would benefit from the real-time estimation of the 3D shapes of these deformable objects from images. This estimation is however challenging due to multiple factors: 1) little information is available within an image to visually detect and distinguish a curvilinear object due to its thin and usually uniform appearance; 2) different 3D shapes may lead to the same visual perception, even in a stereo setting in case portions of the objects lie in an epipolar plane; and 3) the motions and deformations can be large, depending on the stiffness of the object. Additionally, a tracking approach that can consistently track specific points along the object defined by their arclength, such as the extremities or midpoint, would be particularly useful in the aforementioned applications.

To deal with visual ambiguities such as drift along the curve, we propose to texture the object with a coarse pattern of alternating colors and formulate the shape estimation as a deformable 1D template tracking problem. Tracking is expressed as an energy minimization over a set of control points $\mathcal{Q}$ parameterizing a 3D NURBS $\mathcal{C}^{3D}$ modeling the object:

$$\mathcal{C}^{3D}(\mathcal{Q},u) = \sum_{i=1}^{n} R_{i,d}(u)\mathcal{Q}_n, \quad u \in [0,1],$$

where $R_{i,d}$ are the rationale spline basis functions. We make use of the projective invariance properties of NURBS and, in a stereo setup, denote by $\mathcal{C}_1^{2D}$ and $\mathcal{C}_2^{2D}$ the 2D projected splines defined from the projections of $\mathcal{Q}$. The color pattern texturing the thread is represented by a general function associating the curve parameter u to its color $c(u)$:

$$c(u) : u \in [0,1] \longrightarrow S,$$

where S is a color space. For generality, we do not require the two cameras to possess the same color-calibrations, but maintain instead two representations of the texture by using two functions: c_i with $i \in \{1,2\}$ learned from the first images.

Assuming the object's in-extensibility, we propose a novel energy $E = E_{ext} + E_{int}$ based on a texture-sensitive distance map. The first term is a data term enforcing the curvilinear and texture appearances:

$$E_{ext} = \frac{1}{2(K+1)} \sum_{i=1}^{2} \sum_{k=0}^{K} \left(\alpha \mathcal{D}_i^{tub}(\mathcal{C}_i^{2D}(u_k))^2 + \beta \mathcal{D}_i^{tex}(\mathcal{C}_i^{2D}(u_k), u_k)^2 \right),$$

where $\mathcal{D}_i^{tub}(x)$ is a distance map indicating the distance to the closest ridge from position x in image I_i and $\mathcal{D}_i^{tex}(x,u)$ is a texture sensitive distance map indicating the distance to the closest pixel with color $c_i(u)$. $\mathcal{D}^{tex}$ is proposed instead of direct SSD evaluation to increase the convergence radius since the object has a thin structure. $(u_1,\dots,u_K)$ defines a uniform sampling of the parameters and α,β are weights balancing the two terms. The second term enforces the in-extensibilty constraint for a thread with length L_{ref} and maintains an arc-length parameterization:

$$E_{int} = \frac{\gamma}{K} \sum_{k=0}^{K-1} \left(1 - \frac{\int_{u_k}^{u_{k+1}} \|\mathcal{C}^{3D'}(u)\| \mathrm{d}u}{L_{loc}} \right)^2,$$

where $L_{loc} = L_{ref}/K$ and γ is a weighting coefficient. This term also maintains consistency between the parameterizations of the texture c and of the spline $\mathcal{C}^{3D}$, thereby avoiding the re-computation of the arclength parameterization and of the corresponding spline basis functions at each time-step. Optimization is performed using Levenberg-Marquardt.

We demonstrate the benefits of this energy in synthetic and real experiments, using data illustrating the deformation and manipulation of a thread with a da Vinci robot. Usual curve distances [1, 4] are not fully suitable for thread tracking evaluation in the context of robotic manipulation, because they do not properly evaluate the tracking of specific points. We therefore use an *arclength error* based on r specific points from the thread uniquely defined by their arclength parameters $\{v_k | 1 \le k \le r\}$:

$$e_{\text{acl}}^{3D} = \frac{1}{r} \sum_{k=1}^{r} \|\mathcal{C}^{3D}(v_k) - \mathcal{C}_{gt}^{3D}(v_k)\|.$$

In particular, we show that the approach allows for deformable tracking in the absence of normal motion along the curve, a challenging practical situation that occurs frequently in practice when the thread is dragged by one extremity.

[1] Tim Hauke Heibel, Ben Glocker, Martin Groher, Nikos Paragios, Nikos Komodakis, and Nassir Navab. Discrete tracking of parametrized curves. In *Proc. of CVPR*, 2009.

[2] Nicolas Padoy and Gregory D. Hager. Human-machine collaborative surgery using learned models. In *Proc. of ICRA*, 2011.

[3] Jur van den Berg, Stephen Miller, Daniel Duckworth, Humphrey Hu, Andrew Wan, Xiao-Yu Fu, Ken Goldberg, and Pieter Abbeel. Superhuman performance of surgical tasks by robots using iterative learning from human-guided demonstrations. In *Proc. of ICRA*, 2010.

[4] Peng Wang, Terrence Chen, Ying Zhu, Wei Zhang, Shaohua Kevin Zhou, and Dorin Comaniciu. Robust guidewire tracking in fluoroscopy. In *Proc. of CVPR*, 2009.