Model-free Visual Control for Continuum Robot Manipulators via Orientation Adaptation

Mrinal Verghese, Florian Richter, Aaron Gunn, Phil Weissbrod, and Michael Yip
University of California, San Diego, La Jolla CA 92032, USA,
{mtverghe, frichter, ahgunn, pweissbrod, yip}@ucsd.edu

Abstract. We present an orientation adaptive controller to compensate for the effects of highly constrained environments on continuum manipulator actuation. A transformation matrix updated using optimal estimation techniques from optical flow measurements captured by the distal camera is composed with any Jacobian estimation or kinematic model to compensate for these effects. By utilizing domain knowledge to define the structure of this matrix, fewer parameters need to be estimated and a stable controller can be guaranteed. The algorithm is tested on a custom robotic catheter and convergence is shown both empirically and theoretically.

Keywords: Soft Robotics, Control

1 Introduction
Continuum manipulators have shown great promise for Minimally Invasive Surgery in the recent years [1,2]. Their ability to conform to natural anatomy allows greater access within the body without requiring incisions or trauma to the patient. This success has been evidenced by viable commercial products like the Monarch platform™ from Auris Surgical Robotics or the Ion Endoluminal System™ from Intuitive Surgical. Unfortunately, the soft structure of continuum robots bring significant control challenges along with their advantages. The robot has infinite degrees of freedom, and constrained space operation in human anatomy involves unknown contact forces and dynamics that mask measurements of the full robot configuration thus further complicating control [3,4,5,6,7]. At worst, these contacts may result in unstable and dangerous behaviors [8].

1.1 Background
Multiple ideas have been tested against the constrained space control problem. The most generic strategy is to estimate the Jacobian matrix online [8,9], which was shown for setpoint / trajectory regulation as well as hybrid position-force control, and was extended by Wang et al. via an adaptive visual servo controller
2 Verghese et al.

Fig. 1: The plot on the left shows convergence of a target point (in pixel space) with and without our adaptive controller to compensate for unknown contact forces in a heavily constrained environment. The first person endoscopic camera is shown on the right viewing a test fixture and showing the objective.

[12,13]. Efforts have also been made to sense the environment or the continuum manipulator configuration and utilize that information [10]. Tully et al. examined estimating the configuration of a segmented snake robot using an extended Kalman filter on the data from a distal electromagnetic tracker [16]. Other learning methods have also been applied; Melingui et al. implemented an adaptive control algorithm utilizing kernel-based learning for continuum manipulators in free space [17], and Giorelli et al. used a neural network to learn the kinematics of three-tendon actuators [18]. As a whole, actuating continuum manipulators in free space or trivially constrained environments is generally well understood and can be done with moderate accuracy, whereas the problem of reliable control in heavily constrained space such as in many surgical tasks is challenging and open ended.

1.2 Contributions

Our previous work examined a completely model-less control framework to solve this problem by estimating the Jacobian online from a distal camera and tendon tension sensors [8,9,10]. However, without structure to the estimated Jacobian, there is a risk of drift in the matrix or potential singularities. We look to build on our previous work by utilizing domain knowledge about the control problem to derive more structured transformations based on fewer parameters that can be estimated from relatively minimal observations. These parameters are then optimized online to ensure convergence and stability in workspace tasks. We apply this controller on a real robotic catheter to verify our claims and demonstrate the advantages of our methods.
2 Methods

When operating tendon based robotic catheters, any actuation is defined by a linear combination of changes in the lengths of multiple tendons. This actuation can be reasonably modeled in free space by constant curvature models [11]. In constrained space operation, however, unknown contact forces can change the distributed friction along tendon lumens resulting in unpredictable configurations (see Fig 1). For control in the distal camera frame, this results in a control mismatch between control inputs and actual actuation. This poses problems both for human operators who expect consistent control while navigating the body, and autonomous algorithms such as visual servoing, whose feedback controllers rely on accurate models for convergence to setpoints and trajectories.

2.1 Problem Formulation

We consider the problem of model-less control [8] for endoscopes with video feedback, and with setpoints and trajectories defined in pixel space. Actuation is considered in the camera frame where the $z$ axis is the depth, and the $x, y$ axes correspond to the pixel columns and rows respectively. Typically the $z$ axis is controlled by a separate linear insertion joint, and actuation in this direction is not considered in this work. The Jacobian, $J$, evaluated at configuration $q$, relates the actuator velocities to the camera center movement in pixel space:

$$\dot{[x \ y]^T} \approx J [\dot{q}_1 \ \dot{q}_2 \ \ldots \ \dot{q}_n]^T$$

where $\dot{q} = [\dot{q}_1 \ \dot{q}_2 \ \ldots \ \dot{q}_n]^T$ are the joint velocities. The pseudo-inverse of the Jacobian, $J^\dagger$, is used to convert control inputs from the camera frame to changes in joint angles. In constrained conditions, and for that matter, in situations where the kinematic model of the robot is inexact, the model-based Jacobian is inaccurate. Formally we describe this problem as:

$$J.J^\dagger[\dot{x} \ \dot{y}]^T \neq [\dot{x} \ \dot{y}]^T$$

(1)

Fig. 2: Control flow chart that shows the process for model-free learning applied to Jacobian rotation adaptation using visual feedback.
where $J_*$ is the real Jacobian, and the left-hand-side and right-hand-side of the equations represent camera motion vectors in the image frame that are dissimilar.

In this work, corrections are applied by estimating an approximate rotation matrix that, when applied to the Jacobian model $J$, where the rotation is found by minimizing the error between these two vectors in a robust manner (i.e. accounting for measurement and process noise). This process is shown in Fig. 2. Note that these corrections can be combined with other methods of estimation or corrections of the Jacobian, as correcting for rotation alone does not necessarily converge to zero error due to time and history-dependent effects such as viscoelasticity, creep, and hysteresis. Formally, this correction when converged should result in:

$$
\left\| J_* J^\dagger R(\theta) \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} \right\| = \left\| \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} \right\| (2)
$$

where $R(\theta)$ is the rotation matrix correcting the Jacobian is estimated from measurements captured by the endoscopic camera, and $\|\cdot\|$ represents a vector l2-norm.

For an intuitive derivation of the correction, let $R \in \mathbb{R}^{2 \times 2}$ be a general linear correction on the control input. It can be decomposed using singular-value decomposition (SVD):

$$
R = U \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} V^\top (3)
$$

where $U$ and $V$ are unitary matrices and $\sigma_1$ and $\sigma_2$ are the singular values.

If $\sigma_1 \neq \sigma_2$, then the $R$ describes a shear, which is a case left for future work. For this work, it is assumed that $\sigma_1 = \sigma_2$.

This implies that $R$ can be written as a product of a scalar and two unitary matrices, which simplifies to a scalar and a single unitary matrix. The scalar can be seen as the required compensation on the magnitude of the control input to overcome any loss of energy in the system due to frictional and viscoelastic losses from actuation and interaction with the environment. This is near-impossible to estimate due to the non-linear and temporal effects as well as the limited sensing information of where contact occurs within the environment. For purposes of estimation, this scalar is set to 1 and only the unitary matrix is left to estimate. The unitary matrix can be rewritten as a rotation matrix:

$$
R(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} (4)
$$

which only dependent on a single parameter $\theta$.

2.2 Measurement

The image data at point $x, y$ and time $t$ is written as $I(x, y, t)$. Assuming a small motion between frames and time step, a first order approximation for the image
data can be written as:

\[ \mathbf{I}(x + \Delta x_o, y + \Delta y_o, t + \Delta t) = \mathbf{I}(x, y, t) + \frac{\partial \mathbf{I}}{\partial x} \Delta x_o + \frac{\partial \mathbf{I}}{\partial y} \Delta y_o + \frac{\partial \mathbf{I}}{\partial t} \Delta t \]  

(5)

where \( \Delta x_o, \Delta y_o \) is the observed motions from the endoscopic camera. Using the previous notation, the observed motion can be defined by:

\[ \mathbf{J} \mathbf{J}^\dagger [\Delta x \ \Delta y]^{\top} = [\Delta x_o \ \Delta y_o]^{\top} \]  

(6)

where the changes in the discretized control are also assumed to be small. Combining (6) and (2) with \( \Delta x_o, \Delta y_o \) as \( \dot{x}, \dot{y} \) results in:

\[ \mathbf{R}(\theta) [\Delta x_o \ \Delta y_o]^{\top} = [\Delta x \ \Delta y]^{\top} \]  

(7)

Note that the Jacobian, \( \mathbf{J} \mathbf{J}^\dagger \in \mathbb{R}^{2 \times 2} \), is assumed to be a full rank to get this expression. The cases where this does not hold is if the control input does not overcome the losses of energy in the system due to frictional and viscoelastic losses, which is not covered in this work, or if a end-effector collision occurs (which necessitates some additional measurements of contact, as demonstrated in [9]). By simply measuring the angle between the intended control, \( \Delta x, \Delta y \), and the observed motion \( \Delta x_o, \Delta y_o \) yields the angle \( \theta \), required for the correction.

Following the brightness constancy constraint, which says projection of the same point results in the same image data at every frame, the first order approximation for the image data from (5) can be simplified to:

\[ \frac{\partial \mathbf{I}}{\partial x} \Delta x_o + \frac{\partial \mathbf{I}}{\partial y} \Delta y_o + \frac{\partial \mathbf{I}}{\partial t} \Delta t = 0 \]  

(8)

\[ \frac{\partial \mathbf{I}}{\partial x} \dot{x} + \frac{\partial \mathbf{I}}{\partial y} \dot{y} = -\frac{\partial \mathbf{I}}{\partial t} \]  

(9)

where \( \mathbf{v} = [v_x \ v_y]^{\top} \) is the observed optical flow. Therefore, optical flow is directly proportional to the observed motion of continuum manipulator in the camera frame. Optical flow can be measured in a variety of ways; the Lucas-Kanade method utilizing Shi-Tomasi corner detection is used for this work [19,20].

The angle between the observed motion and intended motion can be found by comparing the optical flow result and the control input \( \Delta x, \Delta y \) and is the correction needed. The angle between the two vectors is simply computed:

\[ \theta = \cos^{-1} \left( \frac{[\Delta x \ \Delta y]^{\top} \cdot \mathbf{v}}{\| [\Delta x \ \Delta y]^{\top} \| \| \mathbf{v} \|} \right) \]  

(10)

and \( \theta \) generates the linear correction \( \mathbf{R}(\theta) \).
2.3 Estimation

A measurement of $\theta(t)$ can be taken at every image pair $I(x, y, t)$ and $I(x, y, t-1)$ and is generated from optical flow which is very noisy. The Kalman Filter can be used to reduce noise and have a more accurate measurement. To formulate this as a filtering problem, let $\hat{\theta}(t)$ be the filtered estimate of the correction with variance $\sigma^2_\theta(t)$. The motion model is assumed to have Gaussian noise, so the equation is simply:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + w$$

where $w \sim \mathcal{N}(0, \sigma^2_w)$. Similarly, the measurement model is assumed to have Gaussian noise resulting:

$$\theta(t) = \hat{\theta}(t) + v$$

where $v \sim \mathcal{N}(0, \sigma^2_v)$. Using the Kalman Filter update, the estimate evolves:

$$k(t) = \frac{\sigma^2_\theta(t-1) + \sigma^2_w}{\sigma^2_\theta(t-1) + \sigma^2_w + \sigma^2_v}$$

$$\hat{\theta}(t) = k(t)\hat{\theta}(t-1) + (1-k(t))\theta(t)$$

$$\sigma^2_\theta(t) = (1-k(t))\sigma^2_\theta(t-1) + \sigma^2_v$$

where $k(t)$ is the Kalman gain. This formulation fits the requirements for convergence on the Kalman gain, $k(t)$ [21], so the filter will converge to an Infinite Impulse Response (IIR) filter:

$$\hat{\theta}(t) = \alpha\hat{\theta}(t-1) + (1-\alpha)\theta(t)$$

based on a single parameter $\alpha$. It is easier to tune the single parameter $\alpha$, rather than all the parameters required for the Kalman Filter: $\sigma_\theta(0)$, $\sigma_w$, and $\sigma_v$, so the converged IIR filter is used for estimation. In order to account for the angle wrap around, the final filter used for estimation is:

$$\hat{\theta}(t) = \text{mod} \left( \alpha\hat{\theta}(t-1) + (1-\alpha)\theta(t) + \pi, 2\pi \right) - \pi$$

2.4 Lyapunov Stability

We show under ideal conditions (no hold effects, delay, etc.) the rotation adaptation controller is stable. Let the error be $e = [x_d, y_d] - [x, y]^\top$, where $x_d, y_d$ is some desired position and $x, y$ be current position, and the candidate Lyapunov function is $V = \frac{1}{2}e^\top e$. The resulting derivative of this is:

$$\dot{V} = e^\top \begin{bmatrix} \dot{x}_d \\ \dot{y}_d \end{bmatrix} - e^\top \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix}.$$
In standard position regulation the desired positions would not change, so \( \dot{x}_d, \dot{y}_d \) are both 0, and the error \( e \), would be inputted into the adaptive controller system described in Fig. 2. Therefore, the instantaneous change in output would be:

\[
\begin{bmatrix}
\dot{x} \\
\dot{y}
\end{bmatrix} = J_\star J^\dagger R(\theta) e .
\]  \hspace{1cm} (19)

Substituting this into (18), the resulting derivative of the Lyapunov candidate function is:

\[
\dot{V} = -e^\top J_\star J^\dagger R(\theta) e .
\]  \hspace{1cm} (20)

Through the adaptive controller, the values of \( \theta \) are set such that (2) is satisfied. Therefore the derivative of the Lyapunov candidate function chosen here is:

\[
\dot{V} = -\frac{||J_\star J^\dagger R(\theta)e||}{||e||} e^\top e
\]  \hspace{1cm} (21)

which clearly is always negative except for when the error reaches 0 or when \( J_\star J^\dagger \) is not full rank. Therefore, the system is considered asymptotically stable assuming the control input overcomes the losses of energy in the system, which as stated previously is when \( J_\star J^\dagger \) is not full rank.

3 Experiments

To run the experiments we designed and built a flexible 2.2 mm diameter, meter-long robotic catheter. The robotic catheter is a continuum robot with a flexible backbone comprising five inner channels, one central backbone and four equally spaced (90° separation) around the perimeter. Stainless steel actuation wires/tendons routed through the four radially spaced channels and terminated at the distal end of the catheter provide deflections at the catheter tip. Tendons are terminated on the pulleys in the proximal end, and each pulley is controlled by a separate DC motor. Two tendons, 180° apart represent one antagonistic pair.

Fig. 3: Full view of robotic catheter (left), view of the experimental setup at the distal end (middle), and camera view highlighting the optical flow measurement (right). Features to track for the optical flow were manually added in the environment for this experiment to ensure consistency between tests.
Fig. 4: The three environments that the custom made robotic catheter is tested on. From left to right, the environments are named no bend, one bend, and two bend.

that controls deflection approximately on one plane. Insertion of the catheter is controlled by a separate DC motor that moves the entire assembly on a single linear rail.

All motors are driven by a custom FPGA motor controller system running individual PD loops on each motor [22]. Control was completed utilizing inverse kinematics from the camera frame to the actuators from constant curvature models [11]. The controller interface over ethernet to a desktop running our algorithm in Python. Finally, an endoscopic camera with a total diameter of 2.5 mm with a resolution of 380x400 pixels, a framerate of 30 fps and an integrated LED light ring is attached to the distal end and connected to the desktop through USB. All image processing is done through Python’s OpenCV library.

To test the algorithm in constrained environments, three tortuous paths were made from nylon tubing and metal frames. The environments have zero, one, and two bends and are shown in Fig. 4. The catheter is passed through these environments and then placed in front of an optical marker that was off center in the endoscopic camera’s view, as shown in Fig 1. A simple proportional-controller is used to actuate the catheter and attempt to align the center of the camera with the optical marker. This experimental set up is shown in Fig. 3. Each trial ran for roughly 70 seconds or until it either converged or clearly diverged. The pixel position of the marker and the current value of $\hat{\theta}$ were recorded. When operating a steerable catheter, the operator wants to center a target in the camera frame for insertion, thus pixel distance from the center of the camera feed to the target is an appropriate evaluation criteria. Finally, for each environment the trials were repeated with $\alpha = 1$ (no correction), and $\alpha = .95, .75, .50$.

4 Results

The pixel distance to the target over time for all the experiments are shown in Fig. 5. The full trajectory of the pixel position of the goal for $\alpha = 1, .95$ for the one bend environment is shown in Fig. 1. This visualization highlights the need for orientation adaptation when dealing with the more tortuous environments. The results for the robotic catheter without correction (when $\alpha = 1$) reaffirm the problem. Without any adaptation, the instrument has difficulty converging in increasingly constrained environments, including the no bend environment.
Fig. 5: Pixel distance to target over time for three environments using different values of $\alpha$. From top to bottom the environments are: no bend, one bend, and two bends.
The trajectory visualization in Fig. 1 highlights that when not using orientation adaptation, orbiting around target can occur due to the mismatch in observed and expected actuation and the instrument may never reach said target. The same figure also shows the rapid convergence with the proposed orientation adaptation algorithm. In all tested environments, we were able to see the controller rapidly converge for $\alpha = .95, .75$. With $\alpha = .5$, the optical flow measurements were too noisy preventing $\hat{\theta}$ from converging and leading to instability in the controller.

5 Discussions

Results from our evaluation of adaptation of the Jacobian via a online rotation estimation show the advantages in accuracy and speed of rotation adaptation given different filtering choices. Examining the plots for $\alpha = .95, .75$ shows the trade off between the two values. $\alpha = .75$ converges slightly faster in more constrained environments, at the expense of slightly nosier values for theta and a less smooth trajectory. $\alpha = .95$ yielded a more reliable convergence across all our environments at the expense of taking longer to reach $\hat{\theta}$. Generally this parameter can be tuned based on the quality of the tracked features in the environment and the noise in optical flow.

6 Conclusions and Future Work

Our results showed that estimating only one parameter online still yielded fairly rapid convergence to an accurately mapped Jacobian matrix that would be needed for adaptive and stable control. Such adaptation has significant benefit to both autonomous tasks and obvious benefits to human teleoperators that would observe a correction such that steering commands match the camera directions exactly. In addition, as this single parameter is directly observable, the correction is not prone to drift or artificial singularities. While many completely model free approaches could also converge and generalize very well, integrating domain knowledge about the nature of the changes to the Jacobian helped build a more stable and efficient controller. Furthermore, utilizing a simple linear transform and its SVD decomposition presented in the problem formulation gives good intuition on additional parameters that can be estimated and what they physically represent. A tradeoff is made between the potential model capacity of purely model-free estimation and statistical robustness and stability of structured matrices defined by fewer parameters. This is in part due to the better physical understanding of the perturbations being made to continuum manipulator control to compensate for interactions with constrained environments. Ultimately this stability is essential to seeing this type of control realized in critical surgical applications.
References

1. J. Burgner-Kahrs, D. C. Rucker and H. Choset, Continuum robots for medical applications: A survey, *IEEE Transactions on Robotics*, vol. 31, no. 6, pp. 1261–1280, Dec. 2015.
2. V. Vitiello, S. Lee, T. P. Cundy and G. Yang, Emerging robotic platforms for minimally invasive surgery, *IEEE Reviews in Biomedical Engineering*, vol. 6, pp. 111–126, 2013.
3. R. S. Penning, J. Jung, J. A. Borgstadt, N. J. Ferrier and M. R. Zinn, Towards closed loop control of a continuum robotic manipulator for medical applications, *IEEE International Conference on Robotics and Automation*, pp. 4822–4827, May 2011.
4. R. S. Penning, J. Jung, N. J. Ferrier and M. R. Zinn et al., An evaluation of closed-loop control options for continuum manipulators, *IEEE International Conference on Robotics and Automation*, pp. 5392–5397, May 2012.
5. A. Kapadia and I. D. Walker, "Task-space control of extensible continuum manipulators," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1087–1092, 2011.
6. David B. Camarillo, Christopher R. CarlsonJ. Kenneth Salisbury, "Task-Space control of continuum manipulators with coupled tendon drive, *Experimental Robotics*, vol. 54, pp. 271–280, 2009.
7. V. K. Chitrakaran, A. Behal, D. M. Dawson and I. D. Walker, "Setpoint regulation of continuum robots using a fixed camera," *Robotica*, vol. 25, pp. 581–586, 2007.
8. M. C. Yip and D. B. Camarillo, "Model-Less Feedback Control of Continuum Manipulators in Constrained Environments," *IEEE Transactions on Robotics*, vol. 30, no. 4, pp. 880–889, Aug. 2014.
9. M. C. Yip and D. B. Camarillo, "Model-Less Hybrid Position/Force Control: A Minimalist Approach for Continuum Manipulators in Unknown, Constrained Environments," *IEEE Robotics and Automation Letters*, vol. 1, no. 2, pp. 844–851, July 2016
10. M. C. Yip, J. S. Sganga, D. B. Camarillo, "Autonomous control of continuum robot manipulators for complex cardiac ablation tasks." *Journal of Medical Robotics Research*, vol. 2, no. 1, pp. 1750002, 2017.
11. Webster, R. J., and Jones, B. A., "Design and Kinematic Modeling of Constant Curvature Continuum Robots: A Review," *The International Journal of Robotics Research* vol. 29, no. 13, pp. 1661–1683, 2010.
12. H. Wang, W. Chen, X. Yu, T. Deng, X. Wang and R. Pfeifer, "Visual servo control of cable-driven soft robotic manipulator," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 57–62, 2013.
13. H. Wang, B. Yang, Y. Liu, W. Chen, X. Liang and R. Pfeifer, "Visual Servoing of Soft Robot Manipulator in Constrained Environments With an Adaptive Controller," *IEEE/ASME Transactions on Mechatronics*, vol. 22, no. 1, pp. 41–50, Feb. 2017.
14. Yuen SG, Kesner SB, Vasilyev NV, Del Nido PJ, Robert D, Howe RD, "3D ultrasound-guided motion compensation system for beating heart mitral valve repair." *Med Image Comput Comput Assist Interv*. vol. 11, pt. 1, 2008.
15. S. B. Kesner and R. D. Howe, "Force control of flexible catheter robots for beating heart surgery," *IEEE International Conference on Robotics and Automation*, pp. 1589-1594, 2011.
16. S. Tully, G. Kantor, M. A. Zenati and H. Choset, "Shape estimation for image-guided surgery with a highly articulated snake robot," *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1353-1358, 2011.

17. A. Melingui, J. J. Mvogo Ahanda, O. Lakhal, J. B. Mbede and R. Merzouki, "Adaptive Algorithms for Performance Improvement of a Class of Continuum Manipulators," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 48, no. 9, pp. 1531-1541, Sept. 2018.

18. M. Giorelli, F. Renda, G. Ferri and C. Laschi, "A feed-forward neural network learning the inverse kinetics of a soft cable-driven manipulator moving in three-dimensional space," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 5033-5039, 2013.

19. B. D. Lucas, T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision." *Proceedings of Imaging Understanding Workshop*, pages 121–130, 1981

20. Jianbo Shi, Tomasi, "Good features to track," *(IEEE Conference on Computer Vision and Pattern Recognition)*, pp. 593-600 1994.

21. J. Walrand, A. Dimakis, "Random Processes in Systems – Lecture Notes" https://people.eecs.berkeley.edu/~wlr/226F06/226a.pdf, August 2006.

22. D. Schreiber, D. Shak, A. Norbash, M. Yip, "An Open-Source 7-Axis, Robotic Platform to Enable Dexterous Procedures within CT Scanners" *IEEE/RSJ International Conference on Intelligent Robots and Systems* arXiv:1903.04646, 2019.