Monocular Shape Sensing for Continuum Robot

Thomas Enxu Li, Jimmy Chengnan Shentu, Vicky Chaojun Chen
University of Toronto

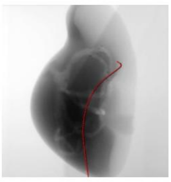
Motivation

- **Continuum Robot** refers to the subcategory of robotic manipulators that do not contain rigid links or identifiable joints. Precise motion control of continuum robots requires real-time and accurate **shape sensing**.
- **Model-based** shape sensing methods are sensitive to unknown external loads, and **sensor-based** methods take up valuable space in the robots and pose challenges to miniaturization.
- Existing **visual-based** shape sensing methods utilize two or more cameras to attain high accuracy, but such conditions may not be achievable in real-world applications.
- We investigate the feasibility of **monocular visual shape estimation** for a continuum robot in terms of **accuracy** and **computation time**.

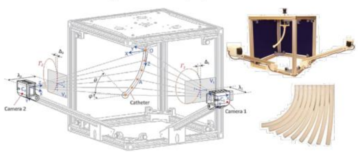
Tendon-driven continuum robot prototype with three extensible sections at different lengths. [1]

Related Work

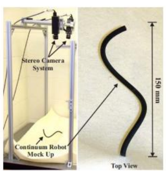
- Burgner et al. achieved a mean error of 0.473 ± 0.353 mm using segmentation and **epipolar geometry analysis** [2].



- Dalvand et al. achieves a maximum measurement error of 0.5 mm for the tip position and length and 0.5 degrees for the bending and orientation angles useing a **stereo vision system** and a 3D reconstruction algorithm [3].

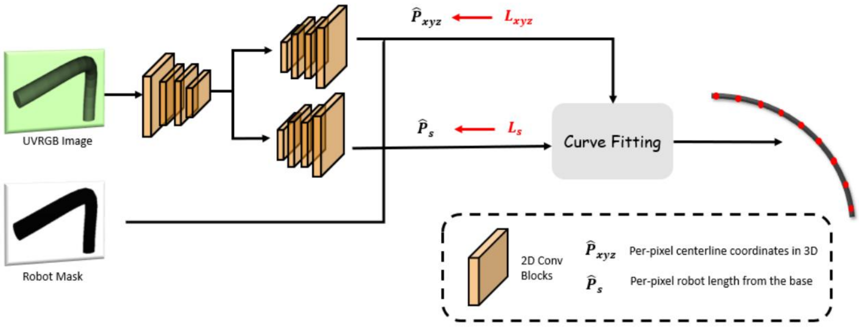


- Croom et al. achieves an average error of 1.53 mm uses a **stereo-vision**-based **self organizing map** [4].



New Technique

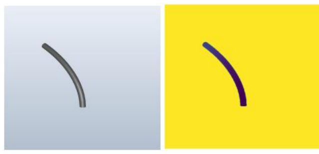
Problem Formulation: Assume we are given an RGB image of the robot, $\boldsymbol{I_{RGB}} \in \mathbb{R}^{H\times W\times 3}$ and a binary occupancy mask of the robot, $\boldsymbol{O} \in \mathbb{B}^{H\times W}$. The goal is to find the position of the robot in 3D, parameterized by the 3D coordinates of M evenly-spaced points on the centerline of the robot, denoted as $\boldsymbol{P_r} \in \mathbb{R}^{M\times 3}$.



Experimental Results

Dataset

- We collected a custom dataset using an existing simulator. The simulated tendon-driven continuum robot is 280 mm in length and 10 mm in radius with a protective sleeve.
- 50,000 randomly sampled robot configurations are rendered with the Visualization Toolkit (VTK), where we save 512 x 512 RGB and depth images along with camera configuration and ground truth robot shape. Texture was added to make the dataset more realistic.
- 80% of the dataset are for training and validation and the remaining 20% are reserved for testing.



Evaluation

- Shape sensing for continuum robots has typically been evaluated in terms of mean error of robot shape (**MERS**) and mean error of tip tracking (**METE**).

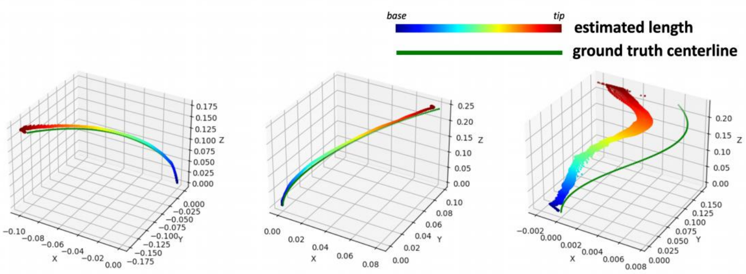
$$MERS = \frac{1}{M}\sum_{j=1}^{M}\left\|\hat{P}_{r,j} - P_{r,j}\right\|_2$$

$$METE = \left\|\hat{P}_{r,M} - P_{r,M}\right\|_2$$

- We also report the runtime of each approach in terms of frames per second (**FPS**).

Architecture	Backbone UN	RN	SN	Decoders D1	D2	D3	D4	Fitting PN	PF	Metrics MERS	METE	FPS
Baseline	✓			✓				✓		16.22	34.47	26.43
	✓				✓				✓	15.74	15.94	26.81
Ablations	✓			✓		✓			✓	15.93	17.45	20.12
		✓		✓		✓			✓	15.40	16.59	11.01
			✓	✓		✓			✓	15.73	17.54	15.94
	✓			✓		✓	✓		✓	1.79	3.48	15.87
Proposed	✓				✓		✓		✓	**1.76**	**3.29**	**20.34**

Ablation study of the proposed components vs baseline. **UN**: Unet [5], **RN**: ResNet [6], **SN**: SalsaNext [7], **D1**: Decoding per-pixel *xyz* coordinates on surface, **D2**: Decoding per-pixel *xyz* coordinates on centerline, **D3**: Decoding per-pixel Δ*xyz* offset from surface to centerline, **D4**: Decoding per-pixel length from robot base, **PN**: PointNet [8], **PF**: Polynomial fitting. **FPS**: measured using Intel(R) Xeon(R) CPU E5-2687W v4 and NVIDIA RTX 2080Ti. Metrics are presented in mm.



References

[1] Neumann, and Burgner-Kahrs. Considerations for follow-the-leader motion of extensible tendon-driven continuum robots. *IEEE International Conference on Robotics and Automation (ICRA)*, 2016.
[2] Burgner, Herrell, and Webster. Toward fluoroscopic shape reconstruction for control of steerable medical devices. ASME Dynamic Systems and Control Conference, 2011.
[3] Dalvand, Nahavandi, and Howe. High speed vision-based 3d reconstruction of continuum robots. IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2016.
[4] Croom, Rucker, Romano, and Webster, Visual Sensing of continuum robot shape using self-organizing maps, IEEE International Conference on Robotics and Automation, 2010.
[5] Ronneberger, Fischer, Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation, in International Conference on Medical Image Computing and Computer-Assisted Intervention, 2015.
[6] He, Zhang, Ren and Sun, Deep Residual Learning for Image Recognition," in IEEE CVPR, 2016.
[7] Cortinhal, Tzelepis, and Aksoy. Salsanext: Fast semantic segmentation of lidar point clouds for autonomous driving, IEEE Intelligent Vehicles Symposium, 2022.
[8] Qi, Su, Mo, and Guibas, Pointnet: Deep learning on point sets for 3d classification and segmentation, in IEEE CVPR, 2017.