Real-Time 3D Reconstruction of Dexterous Continuum Surgical Robots

Tania K. Morimoto*
Department of Mechanical Engineering
Stanford University
Stanford, California 94305
Email: taniakm@stanford.edu

Sean M. Sketch*
Department of Mechanical Engineering
Stanford University
Stanford, California 94305
Email: ssketch@stanford.edu

***Abstract*—Dexterous continuum surgical robots are of increasing interest to researchers given their ability to reach remote areas within the body by moving in highly curved paths. Knowing the configuration of the robot at any time is important for precise control and accurate performance during surgical procedures. Although kinematic models have been developed for some of these robots, these models are only accurate to within several millimeters. The accuracy of these models is of increasing concern as robots are being constructed out of novel materials. In this report, we describe an image processing pipeline for reconstructing the 3D backbone of a concentric tube robot using two orthogonal cameras as the robot moves in free space.**

I. Introduction

CONCENTRIC tube robots are a class of dexterous continuum robots that have been used for a wide variety of applications. Researchers in the medical field have focused on their use as surgical tools, due to their ability to reach remote areas within the body by moving in highly curved paths. They consist of hollow, pre-curved elastic tubes that fit concentrically, each one inside the next. As the tubes are rotated and inserted relative to each other, their curvatures interact to change the robot's overall shape as well as its tip position [1], [2].

Knowing the true configuration of the robot at any given instant is important for precise control schemes and accurate performance during procedures. Several groups have developed kinematic models to aid in the understanding of the robot's position and orientation both in free space and under applied loads [3]. However, these models too must be verified by experimental evaluation and determination of the robot's overall shape. In addition, researchers have started investigating new materials, beyond the standard Nitinol, to make concentric tube robots. It is important to have an accurate method of determining the robot shape and tip position in order to verify that these new materials perform similarly to Nitinol.

Researchers typically mark fiducial points along the backbone of the concentric tube robot to use for tracking and reconstruction [4], [5]. Limitations of fiducial-based methods include reliance on manual point selection, which is lengthy and introduces error, as well as the introduction of error due to the fiducials themselves not being perfect points. An alternative to manual or fiducial-based segmentation of the robot is an automatic reconstruction algorithm that relies on digital image-processing techniques. Camarillo presented such an algorithm [6] for reconstructing a general snake-like robot. In this work, we present a method for 3D reconstruction of a concentric tube robot that uses eigenimage-based background subtraction, morphological centerline extraction, and two-stage voxel carving. We integrate this method with the automatic actuation of a two-tube robot, demonstrating its potential as a verification method for existing kinematic models, as well as the behavior of robots constructed from novel materials.

*These authors contributed equally to this work.

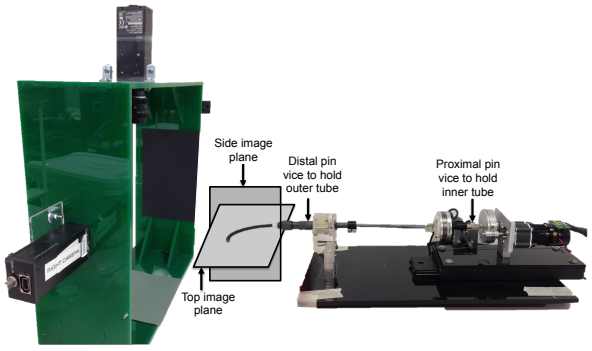

Fig. 1. The experimental apparatus includes the camera mount (left) and the actuation unit (right). The camera mount holds two orthogonal Sony XCD-X710 cameras. The actuation unit consists of two pin vices, distal and proximal, to hold the outer and inner tubes respectively. Two DC motors actuate the rotational and translational degrees of freedom via a capstan drive.

II. Experimental Set-up

The experimental apparatus, including the camera mount and the actuation unit used to drive the concentric tube robot, is shown in Figure 1. The camera mount was built to rigidly hold two Sony XCD-X710 cameras orthogonal to each other and to the actuation unit. The dimensions were designed such that the entire concentric tube robot could be in view at all times by both cameras. The actuation unit used to drive the insertion and rotation of the concentric tubes is similar to the actuation units in [7], [8]. This particular version is for a two-tube concentric tube robot, and there are two pin vices, distal and proximal, to hold the outer and inner tubes respectively. Both the linear and rotational degrees of freedom are actuated via a capstan drive transmission mechanism. All movements are controlled via a MATLAB interface that communicates

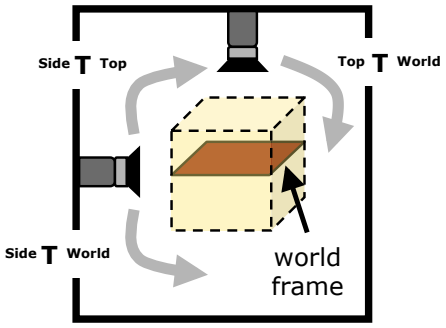

Fig. 2. Camera calibration yielded transformations between the 3 reference frames (side camera, top camera, and world) in our experimental setup. In this schematic, the imaging volume captured by the cameras is represented by the dotted yellow box.

serially with an Arduino Uno which in turn drives the motors.

After the mount and actuation unit were in place, it was important to obtain the intrinsic and extrinsic camera parameters. This was done using MATLAB's Camera Calibration toolbox [9]. The calibration was performed for each camera individually, using 20 distinct images of a standard checkerboard, resulting in the necessary intrinsic and extrinsic camera parameters. The top camera was then used to take a final image with the checkerboard positioned in the desired "world frame" orientation. This provided the homogeneous transformation between the top frame and world frame, ${}^{Top}T^{World}$. A stereo calibration was then performed to obtain the transformation between the two cameras, ${}^{Side}T^{Top}$. The final transformation between the side frame and world frame, ${}^{Side}T^{World}$, was easily obtained using,

$$ {}^{Side}T^{World} = {}^{Side}T^{Top} \cdot {}^{Top}T^{World}. \quad (1) $$

The camera setup, world frame, and relationship between the above mentioned transformation matrices are shown in Figure 2.

III. Algorithms

A. Silhouette Extraction

Silhouette extraction from the raw camera frames was performed in 3 steps: (1) background subtraction, (2) threshold-based binarization, and (3) small-region removal.

To account for camera jitter and variations in environmental lighting, background subtraction relied on a set of eigenbackgrounds [10] instead of a single background image. The eigenbackgrounds (see Fig. 3) were formed by applying the Sirovich and Kirby algorithm to a matrix of frames extracted from a video of the background. The eigenbackgrounds with the five largest eigenvalues were formed into a matrix for projection into the low-dimensional eigenspace. Projecting a camera frame into this eigenspace, then projecting it back into the full image space, preserved frame-specific lighting and positioning while eliminating foreground elements. This background was subtracted from the original camera frame.

The absolute value of the (normalized) difference image (original camera frame minus background) was then binarized using an empirically chosen threshold (0.25). A larger threshold preserved too many undesired parts of the image. A smaller threshold significantly degraded the edges of the tube. With the chosen threshold, the largest connected 1-region was the tube silhouette. It was easily extracted by removing all 1-regions except that containing the most pixels. The 3 steps for silhouette extraction are depicted in Fig. 4.

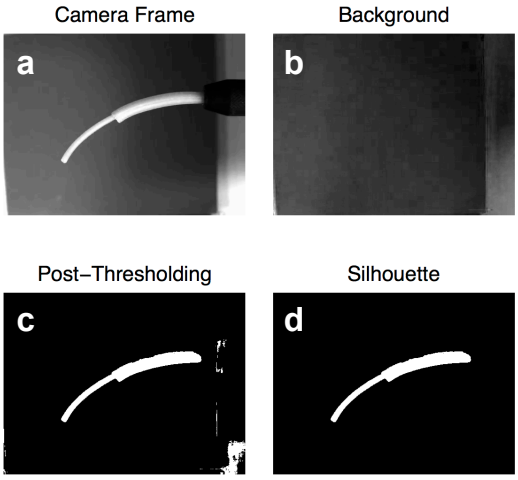

Fig. 4. Beginning with a raw camera frame (a), silhouette extraction was performed in 3 steps: (b) background extraction (and subsequent subtraction) using eigenbackgrounds, (c) binarization using a single threshold, and (d) removal of all except the largest 1-region.

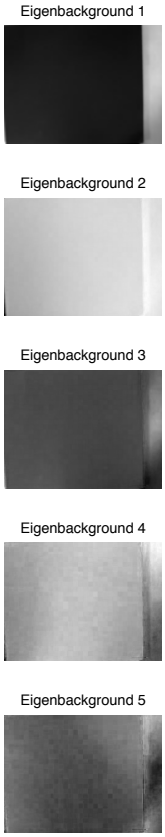

Fig. 3. Five eigenbackgrounds were selected to span the low-dimensional eigenspace used for robust background extraction.

B. Centerline Extraction

Rather than performing voxel carving on the entire silhouette image, the next step of the algorithm was to extract the centerline of the concentric tube robot. This step was performed using a series of morphological image processes, due to the difference in diameter of the inner and outer tubes, which made performing erosion with a single structuring element a challenge. As shown in Figure 5, a large circular structuring element was first used to erode the original silhouette image. This erosion resulted in the centerline of the outer tube, shown in Figure 5a, and it was then subtracted from the original silhouette, to obtain the difference silhouette in Figure 5c. A smaller circular structuring element was then used to perform a second erosion on the difference, in order to obtain the centerline of the inner tube shown in Figure 5d. Small-region removal was performed on both the inner and outer tube centerline images, and the resulting clean centerlines were added together to obtain the final centerline image in Figure 5e.

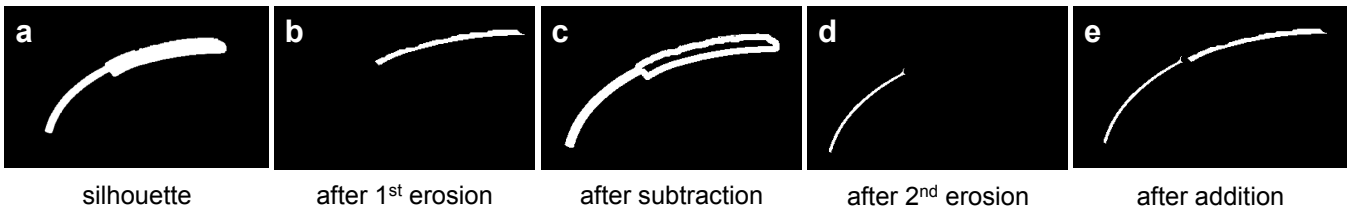

Fig. 5. A series of morphological image processes was performed on the silhouette image (a). A large circular structuring element was used to erode the silhouette, resulting in (b) the centerline of the outer tube. After (c) subtraction of this centerline from the original silhouette, a second erosion with a smaller circular structuring element resulted in (d) the centerline of the inner tube. The two centerlines were then added together to obtain (e) the centerline, or backbone, of the concentric tube robot.

C. Voxel Carving

3D reconstruction of the centerline, or backbone, of the concentric tube robot was done using voxel carving. Based on the dimensions of the silhouettes at a given instant in time, a volume of 100x100x100 voxels was created in the space occupied by the concentric tube robot. These voxels were then projected into the side camera frame using the transformation matrix between the world and side frames. The resulting projected points were compared with the previously determined centerline points. Any of the voxels corresponding to projected points that did not fall within the centerline were "carved" away, resulting in a volume such as that in Figure 6. The remaining voxels were then projected into the top camera frame using the transformation matrix between the world and top frames. Applying the same comparison and carving processes resulted in a volume similar to that in Figure 7. The final reconstruction was visualized by first drawing circles at each point along the backbone (oriented with normals in the direction of the next point along the backbone) and then constructing "patches" between subsequent circles.

IV. Results

The image processing algorithms described above were coded in MATLAB and integrated with live streaming of video from the top and side cameras, along with the automated insertion and rotation of the concentric tube robot. The processing pipeline grabs one frame from each camera whenever a reconstruction is completed. The pipeline has not yet been optimized for speed, and runs between 10 and 15 seconds per frame. To work with this slow frame rate, we inserted and rotated the inner tube in discrete steps separated by 10 seconds. The resulting reconstructions are shown in Fig. 8.

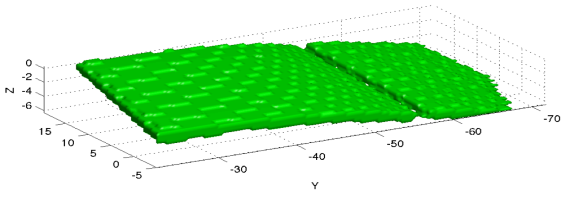

Fig. 6. Result after projecting voxels into side camera frame and carving away any voxels that did not lie within the centerline of the concentric tube robot.

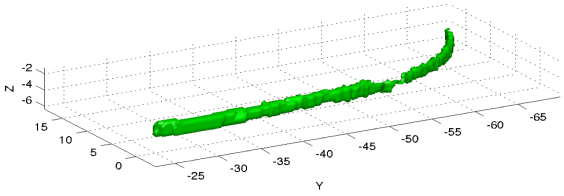

Fig. 7. Result after the second projection of voxels into the top camera frame and carving away any voxels that did not lie within the centerline of the concentric tube robot.

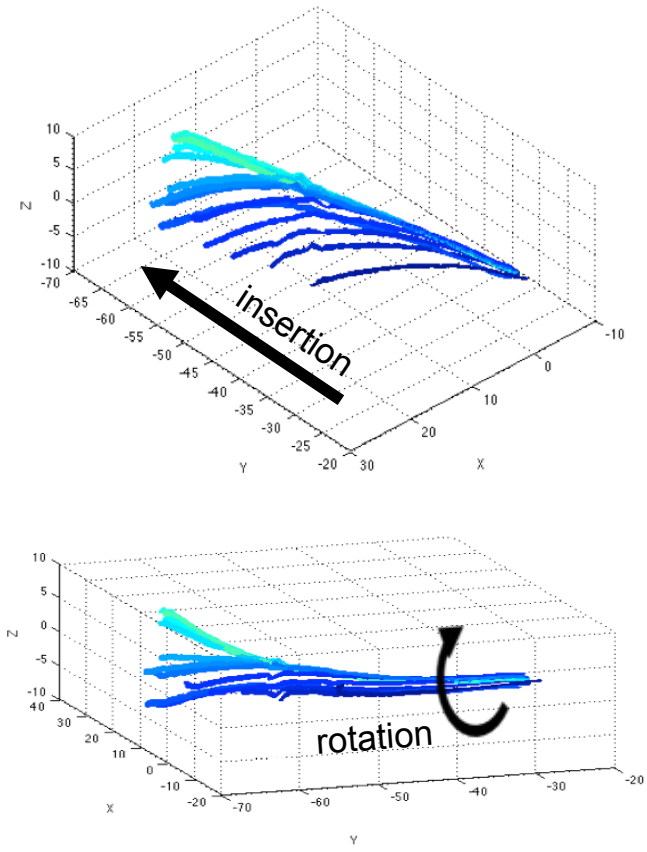

Fig. 8. A two-tube concentric tube robot was reconstructed as it was automatically inserted and rotated. The plotting color progresses from dark to light blue with increasing time. The top panel of the figure more clearly shows the results of insertion. The bottom panel displays the tubes' subsequent upward rotation.

V. Conclusion

We successfully reconstructed the two-tube concentric tube robot in pseudo-real-time (i.e., insertion and rotation speed was slowed to accomodate MATLAB's processing speed). The

reconstruction pipeline will be useful for verifying existing and to-be-developed kinematic models for the robot, as well as any other dexterous continuum surgical robots. Such verification is especially important as the robots begin to be constructed from materials with poorly defined properties (e.g., 3D-printed plastic).

In future work, we aim to first improve visualization of the 3D reconstruction with smoother plotting. Second, we hope to increase the speed and efficiency of the processing pipeline, perhaps via model- or eigenimage- (specifically, "eigentube") based reconstruction. Beyond these natural extensions of the project, we will incorporate a check for consistency between frames, leveraging the temporal nature of the video data. Finally, we could use an Oculus Rift VR headset for real-time display of the reconstructions, providing researchers (and eventually surgeons) a more immersive experience and more realistic representation of the robot.

ACKNOWLEDGMENT

The authors would like to thank Professors Bernd Girod and Gordon Wetzstein and the course assistants for EE 368: Digital Image Processing for their technical support and guidance. They also acknowledge code taken from the MATLAB File Exchange.

REFERENCES

[1] R. Webster, A. Okamura, and N. Cowan, "Toward active cannulas: Miniature snake-like surgical robots," in *IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, 2006, pp. 2857–2863.

[2] P. Sears and P. Dupont, "A steerable needle technology using curved concentric tubes," in *IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, 2006, pp. 2850–2856.

[3] D. C. Rucker, B. A. Jones, and R. J. Webster III, "A Geometrically Exact Model for Externally Loaded Concentric Tube Continuum Robots," *IEEE Transactions on Robotics*, vol. 26, no. 5, pp. 769–780, 2010.

[4] R. Webster, J. Romano, and N. Cowan, "Kinematics and calibration of active cannulas," in *IEEE Int. Conf. Robotics and Automation*, 2008, pp. 3888–3895.

[5] D. C. Rucker, R. J. Webster III, G. S. Chirikjian, and N. J. Cowan, "Equilibrium Conformations of Concentric-Tube Continuum Robots," *International Journal of Robotics Research*, vol. 29, no. 10, pp. 1263–1280, 2010.

[6] D. Camarillo and S. University, *Mechanics and Control of Tendon Driven Continuum Manipulators*. Stanford University, 2008.

[7] E. Butler, R. Hammond-Oakley, S. Chawarski, A. Gosline, P. Codd, T. Anor, J. Madsen, P. Dupont, and J. Lock, "Robotic neuro-endoscope with concentric tube augmentation," in *IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, 2012, pp. 2941–2946.

[8] J. Burgner, D. C. Rucker, H. B. Gilbert, P. J. Swaney, P. T. Russell III, K. D. Weaver, and R. J. Webster III, "A Telerobotic System for Transnasal Surgery," *IEEE/ASME Transactions on Mechatronics*, vol. 19, no. 3, pp. 996–1006, 2014.

[9] J. Y. Bouguet, "Camera calibration toolbox for Matlab," 2008. [Online]. Available: http://www.vision.caltech.edu/bouguetj/calib_doc/.

[10] M. Piccardi, "Background subtraction techniques: a review," in *Systems, man and cybernetics, 2004 IEEE international conference on*, vol. 4. IEEE, 2004, pp. 3099–3104.