

AVA: Automated Viewability Analysis for Ureteroscopic Intrarenal Surgery

Daiwei Lu^a, Yifan Wu^a, Xing Yao^a, Nicholas L. Kavoussi^b, and Ipek Oguz^a

^aDept. of Computer Science, Vanderbilt University, Nashville, TN, USA

^bDept. of Urology, Vanderbilt University Medical Center, Nashville, TN, USA

ABSTRACT

Ureteroscopic intrarenal surgery comprises the passage of a flexible ureteroscope through the ureter into the kidney and is commonly used for the treatment of kidney stones or upper tract urothelial carcinoma (UTUC). Flexible ureteroscopes (fURS) are limited by their visualization ability and fragility, which can cause missed regions during the procedure in hard-to-visualize locations and/or due to scope breakage. This contributes to a high recurrence rate for both kidney stone and UTUC patients. We introduce an automated patient-specific analysis for determining viewability in the renal collecting system using pre-operative CT scans.

Keywords: Viewability, Simulation, Urology, Segmentation

1. INTRODUCTION

Ureteroscopic intrarenal surgery enables a minimally invasive examination and treatment for diseases of the upper urinary tract such as kidney stones and upper tract urothelial carcinoma (UTUC). The operation employs a flexible ureteroscope (fURS): an endoscope composed of a long and flexible steerable tube with a camera at the tip. It is the standard of care for stone treatment and is considered a potential option for less radical UTUC treatment than removal of the whole kidney. However, both diseases face high recurrence or persistence rates. Kidney stones have a 30% risk of repeat surgery¹ and UTUC has a persistence rate of 48-60% after endoscopic ablation due to missed tumors or incomplete treatment.²

The physical limitations of scopes prevent some regions from being visualized during ureteroscopic surgery. This can lead to scope breakages as surgeons force scopes beyond their limits, which impedes the surgery and increases surgical costs.³ Informing a surgeon pre-operatively of which intrarenal regions are viewable with a fURS could reduce the incomplete treatment rates of stones and UTUC, as well as scope breakages. Instead, surgeons could justify percutaneous renal surgery, which is more morbid than ureteroscopic surgery but can enable access to some intrarenal regions not easily viewable using a fURS.⁴

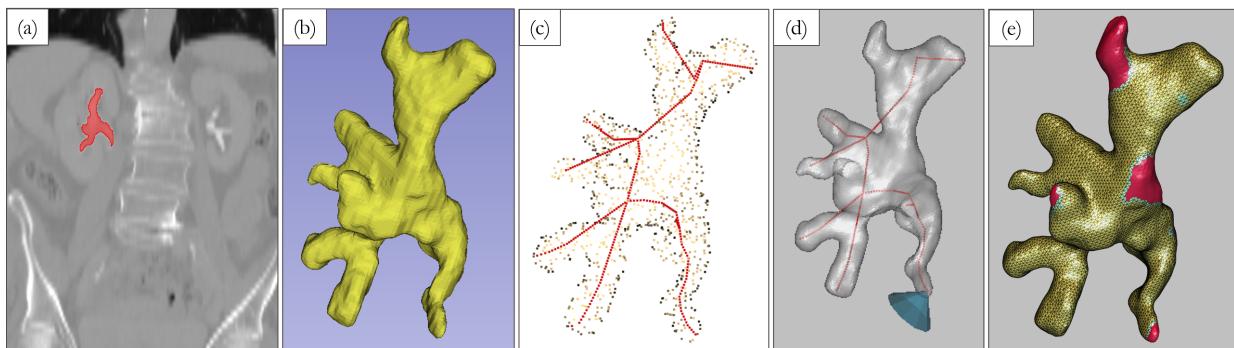


Figure 1. (a) Pre-operative CT imaging. (b) Slicer3D rendering of renal collecting system. (c) Collecting system with skeleton points generated via wave propagation. (d) Camera simulation along skeleton; cone is shown in blue. (e) Viewability analysis output. Yellow: viewable, red: unviewable.

Pre-operative cross-sectional imaging (e.g., computerized tomography (CT) imaging) is the current standard of care for surgical planning. There are no current methods of determining viewability from pre-operative imaging. Thus viewability prediction is currently an unmet need with the potential to improve patient outcomes after ureteroscopic renal surgery.

We introduce a prototype system (Fig. 1) to predict viewability from pre-operative CT images with minimal computational time and little user input. We train a 3D U-Net model to segment the renal collecting system from the pre-operative CT image of a patient. We then compute its skeleton to represent the branching structure. We use this skeleton for simulated traversal of the patient collecting system, starting from the ureter. The output is a map of the collecting system with predicted viewable vs. unviewable locations.

2. RELATED WORKS

Naito et al. analyzed the factors impacting reachability of lesions during transbronchial biopsies.⁵ They considered 4 indices for significance in determining whether lesions were reachable: local curvature at each branch, plane rotation of the scope, anatomical radius, and global relative angle of the scope position. They found that all but the global relative angle were significant. Critically, this analysis only considered single point reachability and thus was only evaluated on 15 unreachable and 26 reachable points. We note that plane rotation is less impactful for reachability during ureteroscopy due to the intrarenal geometry.

Reachability in robotics is a well explored field of research. Kinematic equations have been developed for modeling and controlling the actuated tips of the endoscopes or similar tendon/cable continuum robots.^{6–8} However, the reachability problem in the context of robotics generally aims to solve robot positioning in free space,⁹ and would require robust simulation to accurately determine reachability across an entire anatomical region. In contrast, we aim to estimate reachability by simulating an agent with simple constraints and significantly less computation than a full-scale endoscope model.

3. METHODS

3.1 Data

We use 17 pre-operative Siemens CT scans (Fig. 1a) obtained for pre-operative planning or disease surveillance in our experiments. This included 12 patients with UTUC, 3 with kidney stones, and 2 with negative surveillance imaging. The kidneys and collecting systems are manually labeled by a student and verified by a board certified urologist using ITK-SNAP’s Active Contour tool (itksnap.org). 4 of the UTUC subjects also have surgical treatment videos used for the viewability analysis verification (Sec. 3.6).

3.2 Endoscope Movement Measurement

To obtain parameters to model endoscopic behavior, we mark the movement range of a fURS (LithoVue, Boston Scientific) in a rigid 3D printed phantom (Fig. 2). The angular range of movement is then measured from the center-line of each pole.

We measure scope tip movement both when constrained in a narrow renal infundibulum (i.e., branch of the collecting system) and inside the wider renal pelvis. We term this the **local flexion ability** of the scope tip (i.e., the ability to flex the scope within the constraints of different parts of the renal anatomy). Additionally we measure the maneuverability of the scope to reach each calyx. We term this the **global flexion angle** of the scope shaft, which captures how acute of an angle the scope can flex within the phantom’s overall structure.

3.3 CT Segmentation

For CT segmentation pre-processing, CT scans are first re-sampled to reduce the variability in the CNN input and alleviate low out-of-plane resolution. The intensity of the scans is then clipped and normalized. Image augmentation consists of random intensity shift, random cropping of samples with a balanced foreground-background ratio, random rotation, random scaling, and a Gaussian smoothing.

Next, we train a 3D U-Net model¹⁰ implemented in MONAI (monai.io). The input of the UNet model is a patch of $128 \times 128 \times 128$ voxels due to computational constraints. For prediction, we use sliding window inference

with an overlap ratio of 0.5 and use the mean of overlapping predictions for the final segmentation result. The model output is a segmentation of the entire kidneys. A 6-fold cross validation experiment is run with 11 train/3 validation/3 testing subjects per fold.

In post-processing, the kidney segmentation is dilated by a $5 \times 9 \times 9$ vox structuring element, ensuring the segmentation captures the whole collecting system. Then, we mask the original CT scan using this dilated kidney segmentation. Finally, we perform a three-class Otsu thresholding¹¹ and select the highest intensity label to isolate the collecting system, which is generally highlighted in delayed-phase CT.

3.4 Skeletonization

The binary segmentation of the renal pelvis is converted into a triangle surface mesh using the marching cubes algorithm¹² and smoothed with a Gaussian kernel, as implemented in ITK-SNAP. We extract the skeleton of this surface based on a wave propagation algorithm* from the computer graphics literature.¹³ Wave propagation calculates the number of steps to reach each vertex from a centerpoint and treats vertices with equal step sizes as a layer which is then contracted to form a center point. The step size parameter adjusts how many vertices are contracted at each layer and wave count adjusts how many iterations are run which are then averaged. This algorithm was selected because of its good performance on tubular structures similar to the surface meshes of the renal collecting system.

3.5 Viewability

All rendering and viewability calculations are done through the Visualization ToolKit (VTK).¹⁴ The user is prompted through a GUI to select a point on the mesh where the lowest point of the ureter is, which is the entry point for the simulated scope.

Beginning from the closest skeleton point to this ureter entry point, we construct a breadth-first search tree of the directed skeleton edges. At each edge, the angle from the previous edge is considered. If it exceeds the global flexion angle, then that edge and its successors are no longer considered. The result of this is an ordered traversal tree representing everywhere the endoscope can travel along the skeleton.

The tree is traversed edge by edge inward from the ureter. Along each edge, camera positions are sampled at a constant step size. Camera directions are then generated to be evenly distributed along a spherical sector using a Fibonacci lattice. The sector size is determined by the local flexion angle. Along each camera direction,

*<https://github.com/navis-org/skeletor>



Figure 2. Marking FURS movement range in a 3D printed PLA phantom of the renal collecting system. Red dots show the farthest bound of tip flexion within each branch. **Left**, the global flexion of the scope is shown by bending to the upper-most circle. **Center**, the local flexion ability of the scope in the upper pole calyx. **Right**, viewability computation. The camera is placed along the ureter, and a cone (blue) originates at the camera. The surface mesh cells intersecting with the cone are marked in yellow. Teal triangles mark the boundary between seen and unseen.

a cone is then placed to represent the camera view with the scope tip lying on the skeleton edge. The cone angle is determined by the camera FOV determined from the scope. Finally, the surface mesh is intersected with each cone and the overlapping VTK cells are marked as “seen”. An example of this step can be seen in Fig. 2.

The process is repeated for each sampled point along the constructed traversal tree and combined through an inclusive-OR operation to generate the final viewability map.

3.6 Viewability Analysis Verification

For the 4 UTUC subjects where surgical video are available, the recordings are cropped to the visualization stage of the operation where the surgeon explores the environment before any ablation occurs.

We then create a system allowing the user to place points along a surface mesh extracted from the preoperative CT imaging segmentation result. An expert surgeon (fellowship trained endourologist, case volume of >100 FURS cases per year) is tasked with watching the associated surgical videos and marking the unseen regions (shown in Fig. 3). These cells (triangles) on the surface mesh are marked as unviewable. The unmarked cells of the surface mesh are treated as viewable cells. These are stored in a boolean array for each cell in the mesh and used as the ‘ground truth’ for the automated viewability analysis.

We report the intersection over union (IoU) and the Dice similarity coefficient (DSC). These measures were calculated by comparing the boolean array values in the ground truth and our AVA predictions.

4. RESULTS

4.1 Analysis

The scope is modeled using three measures: camera FOV, local flexion angle at the tip, and global flexion angle along the scope shaft. Camera FOV was determined to be 87° degrees, which determines the width of our simulated cone. The local flexion angle at the tip was approximated at 20° when restricted in a branch and 45° when in the pelvis. This determines how far off axis the camera can move at any point in the collecting system.

The scope global flexion angle was measured at 120°. This determines the angle between skeleton edges we consider to be traversable in our simulation. We note that this is much smaller than the 270° flexion angle specified in the scope manual due to the narrowness inside of the collecting system. While the scope can flex to 270° without any hindrances, we found that the tissue walls frequently restricted and prevented such large flexion from occurring.

4.2 CT Segmentation

Using six-fold cross validation, our 3D UNet model achieved an average DSC score of 0.853 ± 0.084 for collecting system segmentation. Visual inspection of these segmentation results shows that the branching and continuity of the collecting system are preserved even when the boundary placement is locally imperfect. Therefore, the skeletonization function is able to successfully generate the wave function even in the presence of over/under-segmented boundaries.

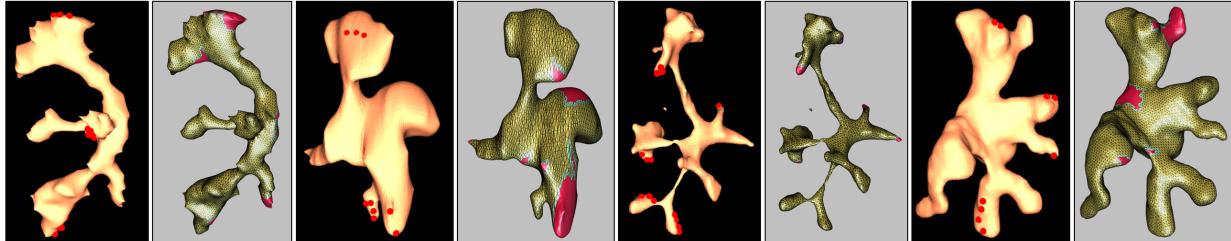


Figure 3. Expert labels (black background) and the AVA viewability predictions (gray background) in 4 subjects. Red dots are expert labels marked in unviewable regions. Yellow: viewable, red: not viewable. The system appears to perform better in narrow regions; e.g., in the second and fourth columns, parts of the renal pelvis are incorrectly marked as unviewable.

4.3 Viewability Analysis Verification

The locations marked as viewable/unviewable by the surgeon are displayed in Fig. 3 (black panels). Note that the unviewable regions are all distributed towards the end of each calyx as opposed to along any of the renal pelvises.

AVA viewability predictions are also shown in Fig. 3 (gray panels). We observe overall consistency between several of the extreme lower and upper pole calyces of the expert labels and AVA predicted regions. The mean IoU was 0.854 and mean Dice score was 0.921 across all 4 renal collecting systems.

However, there are also some dissimilarities. In the second subject, a portion of the lower calyx is erroneously marked as unreachable. Similarly in both the second and fourth subjects, there are regions in the renal pelvis (the central chamber) falsely predicted to be unreachable. These discrepancies occur where both meshes are wide and bulbous. This shape may negatively impact the paths generated during skeletonization as the shape becomes difficult to represent as a single axis.

The system also showed sensitivity to the generated cone depth. A deeper cone would reduce the unviewable regions, and would likely correct the false negative errors in the large central chamber viewability. However, it would also likely cause false positives in areas *behind* narrow branches. This is because we currently mark not just the closest cell intersecting the cone, but rather all intersecting cells, as viewable. Replacing this with ray-casting to enable occlusion handling remains as future work.

5. CONCLUSION

To the best of our knowledge, we present the first ever ureteroscopic viewability analysis for the intrarenal collecting system. We verified our system, AVA, against an experienced surgeon’s annotations paired with surgical videos from patients. Our system parameterizes flexible ureteroscope behavior with a series of measured parameters which can be adapted for different scopes.

In a verification study, AVA predictions showed agreement with surgeon-labeled viewability; the accuracy was higher in narrower areas of the collecting system model. This limitation will be addressed in future work by adjusting the skeletonization algorithm to generate more paths within wider renal pelvises. The small study size may not capture all of the variability of renal collecting systems. While plane rotation is not considered within this study, more data would allow us to determine whether it does affect viewability. Validation will be extended by incorporating more data and additional experts in future work.

Despite these limitations, our proposed AVA system shows validity in viewability prediction. It shows potential for preoperative planning and might in future help reduce the persistence of kidney stones and UTUC.

ACKNOWLEDGMENTS

This work was supported in part by the National Institutes of Health (R21DK133742-01), the Vanderbilt Institute for Surgery and Engineering (VISE) Physician in Residence Program, and the Ruth L. Kirschstein National Research Service Fellowship (3T32EB021937-08S1). This study was conducted in part using the Advanced Computing Center for Research and Education at Vanderbilt University.

REFERENCES

- [1] Khan, S. R., Pearle, M. S., Robertson, W. G., Gambaro, G., Canales, B. K., Doizi, S., Traxer, O., and Tiselius, H.-G., “Kidney stones,” *Nature reviews Disease primers* **2**(1), 1–23 (2016).
- [2] Browne, B. M., Stensland, K. D., Moynihan, M. J., and Canes, D., “An analysis of staging and treatment trends for UTUC in the National Cancer Database,” *Clinical Genitourinary Cancer* **16**(4), e743–e750 (2018).
- [3] Martin, C. J., McAdams, S. B., Abdul-Muhsin, H., Lim, V. M., Nunez-Nateras, R., Tyson, M. D., and Humphreys, M. R., “The economic implications of a reusable flexible digital ureteroscope: a cost-benefit analysis,” *The Journal of urology* **197**(3), 730–735 (2017).
- [4] Rodríguez, D. and Sacco, D. E., “Minimally invasive surgical treatment for kidney stone disease,” *Advances in chronic kidney disease* **22**(4), 266–272 (2015).

- [5] Naito, M., Masaki, F., Lisk, R., Tsukada, H., and Hata, N., “Predicting reachability to peripheral lesions in transbronchial biopsies using ct-derived geometrical attributes of the bronchial route,” *International Journal of Computer Assisted Radiology and Surgery* **18**(2), 247–255 (2023).
- [6] Burgner-Kahrs, J., Rucker, D. C., and Choset, H., “Continuum robots for medical applications: A survey,” *IEEE Transactions on Robotics* **31**(6), 1261–1280 (2015).
- [7] Lipkin, H., Munnae, J., Zhou, G., and Daley, W., “Endoscope kinematics,” in [*12th IFToMM World Congress, Besanon*], (2007).
- [8] Webster III, R. J. and Jones, B. A., “Design and kinematic modeling of constant curvature continuum robots: A review,” *The International Journal of Robotics Research* **29**(13), 1661–1683 (2010).
- [9] Ying, Z. and Iyengar, S. S., “Robot reachability problem: A nonlinear optimization approach,” *Journal of Intelligent and Robotic Systems* **12**, 87–100 (1995).
- [10] Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., and Ronneberger, O., “3D U-Net: learning dense volumetric segmentation from sparse annotation,” in [*MICCAI*], 424–432, Springer (2016).
- [11] Otsu, N., “A threshold selection method from gray-level histograms,” *IEEE transactions on systems, man, and cybernetics* **9**(1), 62–66 (1979).
- [12] Lorensen, W. E. and Cline, H. E., “Marching cubes: A high resolution 3d surface construction algorithm,” *ACM siggraph computer graphics* **21**(4), 163–169 (1987).
- [13] Schreiber, Y. and Sharir, M., “An optimal-time algorithm for shortest paths on a convex polytope in 3D,” in [*Proceedings of the 22nd annual symposium on Computational geometry*], 30–39, Computational Geometry (2006).
- [14] Schroeder, W., Martin, K., and Lorensen, B., [*The Visualization Toolkit (4th ed.)*], Kitware (2006).