Through the Eustachian Tube and Beyond: A New Miniature Robotic Endoscope to See Into the Middle Ear


Abstract—This letter presents a novel miniature robotic endoscope that is small enough to pass through the Eustachian tube and provides visualization of the middle ear (ME). The device features a miniature bending tip previously conceived of as a small-scale robotic wrist that has been adapted to carry and aim a small chip-tip camera and fiber-optic light sources. The motivation for trans-Eustachian tube ME inspection is to provide a natural-orifice-based route to the ME that does not require cutting or lifting the eardrum, as is currently required. In this letter, we first perform an analysis of the ME anatomy and use a computational design optimization platform to derive the kinematic requirements for endoscopic inspection of the ME through the Eustachian tube. Based on these requirements, we fabricate the proposed device and use it to demonstrate the feasibility of ME inspection in an anthropomorphic model, i.e., a 3-D printed ME phantom generated from patient image data. We show that our prototype provides >74% visibility coverage of the sinus tympani, a region of the ME crucial for diagnosis, compared to an average of only 6.9% using a straight nonarticulated endoscope through the Eustachian tube.

Index Terms—Medical robots and systems, steerable catheters/needles, surgical robotics.

I. INTRODUCTION

O

VER the last decade, many robotic systems have been designed for minimally invasive medical applications [1].

These devices are typically deployed within the human body through small incisions or natural orifices, and are used in a variety of applications, e.g. to deliver treatment or conduct visual diagnosis. The natural orifice paradigm in particular is beneficial because it does not require damage to healthy tissue that might be in the way for the surgeon between the body entry point and the diagnostic/interventional site. One natural orifice that has not previously been used for a robotic natural orifice procedure is the Eustachian tube, which connects the nasopharynx to the middle ear (see path labeled ET in Fig. 1).

The motivation for inspection and therapy delivery in the middle ear (ME) comes from the high incidence of ME disease. ME disease is made up of a number of medical conditions that affect the internal structures of the ear, and are estimated to affect approximately 1 in 20 adults worldwide [3]–[5]. One condition of particular interest is cholesteatoma, a benign tumor that erodes hearing structures and causes hearing loss and recurrent infections. Treatment requires surgical removal, but cholesteatoma is challenging to diagnose since it occurs behind the eardrum and externally visible symptoms are non-specific [6]. As a result, cholesteatoma is often missed in primary care and diagnosed only later in the disease progress, typically after permanent damage to the internal hearing structures has occurred.

Fig. 1. An axial computed tomography (CT) scan illustrating the traditional surgical approach to the middle ear through the external auditory canal (EAC), in which the eardrum (commonly referred to as tympanic membrane or TM in the medical literature) must be surgically incised. Also illustrated is the Eustachian tube (ET) approach. Adapted from [2]. The dashed line indicates the position of the axial CT slice.
To accurately diagnose cholesteatoma and other ME diseases, direct visualization is often necessary, which currently requires surgery. The eardrum is cut or lifted and either an endoscope or microscope is used to visualize internal ear structures [10] through the external auditory canal (path labeled EAC in Fig. 1). Importantly, many cholesteatoma patients (nearly 60% according to a recent study [11]) undergo a follow-up surgery approximately one year after the initial tumor removal. This surgery is required to inspect the ME and determine whether the cholesteatoma has recurred. Recurrence happens only about 25% of the time, meaning that many patients have a second surgery, but do not have recurrence [11]–[13]. The non-invasive approach through the Eustachian tube that we propose in this letter could save these patients a second surgery, and more broadly help physicians diagnose and treat a variety of ME diseases less invasively.

The feasibility of passing an endoscope through the Eustachian tube has been previously established by physicians [14]–[17], but inspection usefulness has been limited by (1) lack of a bending tip on the small endoscopes used, and (2) the relatively low resolution provided by the flexible fiberscopes employed. Fortunately, two advancements have recently been made that address both these limitations. First, the latest generation of digital chip-based cameras provide approximately an order of magnitude higher resolution than similarly sized fiberscopes, i.e. over 100,000 pixels as in [18]. Next, an asymmetric flexure wrist design has recently been proposed in the robotics literature that enables a bendable tip to be created in a small diameter, with a large open lumen [19], [20]. Small diameter is required, because it has been established that devices must be smaller than 2 mm (and ideally as small as possible) to pass through the Eustachian tube [21]. In this letter, we combine the bending tip design of [19], [20] with a miniature chip-tip camera to create an endoscope suitable for ME inspection. This is the first report of an endoscope with controllable tip articulation designed to pass through the Eustachian tube.

Our study is structured as follows: we first derive general kinematic requirements for endoscope operation inside the ME. This analysis is performed by simulating the endoscopic inspection on three-dimensional models of the ME extracted from the CT scans of real patients. Using these models, we determine appropriate ranges for the endoscope diameter, length and required curvature. Later in the letter, we use these requirements to guide the design of the proposed device. Finally, we fabricate a prototype of the ME endoscope and demonstrate the feasibility of endoscopic ME inspection in an anthropomorphic phantom, matching the geometry of the human middle ear. Results of this study are discussed in the light of recent studies that have theorized the feasibility of ME endoscopy through the ET [2], [22].

II. KINEMATIC REQUIREMENTS

The anatomy of the ME is illustrated in Fig. 2. It consists of a main chamber, whose average volume in adults is 6.5 cm³ [25], and a number of smaller peripheral cavities. The main chamber contains the ossicles, i.e. a chain of bony structures responsible for the transmission of sound from the eardrum to the internal hearing organs (cochlea). The procedure we wish to enable consists of performing a visual inspection of the ME using an endoscope that is passed through the ET. We are particularly interested in enabling the inspection of the regions where disease is more likely to develop. Cholesteatoma, particularly recurrent cholesteatoma, is frequently found in the sinus tympani [26], a cavity that branches out of the main ME chamber (highlighted in Fig. 2). Because of its anatomic location, the sinus tympani (ST) cannot be reached from the ET via a straight-line path [2]. This is especially evident in Fig. 3(a), which shows the endoscope direction of insertion into the ME and the location of the ST. We hypothesize that an endoscope equipped with a steerable tip will enable a level of visual coverage that exceeds the performance of traditional straight scopes.

To determine the general kinematic requirements for endoscope inspection in the ME, we developed a software platform to simulate this procedure. The software uses 3D anatomical models of the ME generated from CT scans of real patients, and implements the ray-casting algorithm described in [2] to calculate the visible surface area of the ST during each simulated inspection. Repeated endoscope inspections were simulated to determine a suitable range for the endoscope length and curvature.

A. Geometric Computer Model of the Middle Ear

Models of ME geometry were created according to the methods described in [2]. These methods involve the use of statistical atlases, generated from high-resolution micro CT scans, to assist in the segmentation of ME structures from individual patient standard CT scans. For the following analyses, a total of six ME models were generated from patient CT scans.

B. Simulated Endoscope Insertion in the Middle Ear

Repeated insertion paths were generated as follows: a path \( \gamma_i \) consists of a straight section (represented by blue dots in...
Fig. 3. Simulated endoscopy of the middle ear. (a) shows an example of endoscope insertion: the device shown here has a diameter of 1.40 mm and has been bent in such a way to point towards the sinus tympani (ST). (b) illustrates how insertion paths $\gamma_i$ are generated. The endoscope enters the ME following a straight path $\hat{u}_i$ which coincides with the orientation of the Eustachian tube (measured in the CT scan). After an initial straight motion of arbitrary length, the endoscope is bent along a constant curvature arc towards a target point. Target points are generated using a spherical grid in the proximity of the target anatomy. Insertion paths found to collide with the ossicles (like $\gamma_2$ and $\gamma_3$ in this example) are automatically detected and discarded. The endoscope final orientation $\hat{u}_f$ defines the overall bending $\theta$.

Fig. 3(b)) when it is initially inserted into the ME, and a constant curvature section when it is bent towards the target point (dashed line in Fig. 3(b)). In all the simulation trials, the endoscope was navigated from the ET to the entrance of the ST cavity (selected manually in CT scan), as illustrated in Fig. 3(a). To simulate different lengths and curvatures for the endoscope steerable section, we generated a high number of insertion paths—each time incrementing the length of the straight section by 0.10 mm—leaving the length and curvature of the curved section free to vary (Fig. 3). Note that in these simulations we assumed an outside diameter of 1.4 mm for the endoscope. In Section III we will discuss the feasibility of creating a prototype at this diameter. We note that the optimal diameter to be used clinically is still an open question, but that non-steerable endoscopes with diameters as large as 2 mm have been passed through the eustachian tube in patients previously [14]–[17].

C. Calculation of Kinematic Requirements

1) Endoscope Curvature: For each of the six ME models, we determined the insertion path that maximizes visual coverage of the ST and calculated the associated bending angle [i.e. $\theta$, as shown in the call-out in Fig. 3(b)] and length of the curved section. Results are summarized in Table I. This calculation was performed using the ray-casting technique described in [2], and assuming an endoscope field of view (FoV) of 90° (i.e. the same FoV available in many commercially-available endoscopic imaging systems, e.g. [17], [18]). Bending angles range between 47.8° and 74.8°, therefore any design offering at least 74.8° of angular deflection will be suitable for all the ME specimens considered here.

2) Endoscope Tip Length: The curved section lengths reported in Table I provide an indication on the overall length of the endoscope steerable tip. We conducted another round of simulated insertion trials, this time limiting the maximum length of the endoscope curved section to 13.45 mm. This value was determined as the mean between the upper and lower values observed earlier (i.e. 15.57 and 11.33 mm). Table II reports the ST visualization that was achieved under this condition. These results present negligible differences with respect to the optimal visualization values calculated earlier, therefore suggesting that 13.45 mm is an adequate length for the endoscope steerable tip.

III. ENDOSCOPE DESIGN AND FABRICATION

Based on the requirements derived in the previous section, here we present the design of the proposed ME endoscope. The device concept is shown in Fig. 4: it consists of a flexible Nitinol sheath with a chip-tip camera mounted at its tip. The asymmetric notches on the body of the device enable unidirectional bending ($\theta$), which combined with translation ($z$) and rotation ($\varphi$) enable the device to reach virtually any point in 3D. In the following, we first describe the components of the device, and then report on its fabrication.

A. Chip-Tip Camera

The chip-tip camera is the minnieScope-XS (Enable Inc., Redwood City, CA), which has a diameter of 1.4 mm and a length of 2 mm. This CMOS-based videoscope offers a
resolution of 400 × 400, a field of view of 90°, and a frame rate of 30 frames per second. Four optical fibers are integrated on the tip to provide illumination. The camera is connected to an external video processing unit through a flexible cable (diameter: 1.32 mm) that contains both electrical connections and the optical fibers used for illumination.

B. Bending Mechanism

The distal tip of the Nitinol tube is notched, creating a steerable section that can be controllably curved using a thin tendon. This mechanism was originally described in [19], and is particularly suited for our current purposes since (1) it provides a large open lumen for the camera and optical fibers used for illumination, (2) it can be easily miniaturized below 2 mm (as small as 0.46 mm was demonstrated in [20]), and (3) it has been shown to enable tighter radii of curvature than other similar designs, e.g. [27]–[29].

Bending control is enabled by a model that maps the tendon displacement, Δl, to the resulting bending angle. This model is described in [19], [20], but is briefly reviewed here for completeness. For a single cutout, such a mapping can be found using the tube inner radius ri and the chord function (refer to Fig. 5),

\[
\Delta l = h - t
\]

\[
\Delta l = h - 2 \left( \frac{1}{\kappa} - r_i \right) \sin \left( \frac{\theta}{2} \right)
\]

where Δl varies with the actuated length of the tendon within the cutout, t. θ is the bending angle of a single section and κ is the curvature about the wrist center axis (red dashed line in Fig. 5) and is given by

\[
\kappa = \frac{\theta}{h - \bar{y} \theta}
\]

Here, \( \bar{y} \) represents the position of the neutral bending plane, i.e. the region of the tube that experiences no longitudinal strain during bending. Such a location depends on the depth of the cut \( w \), as well as the inner and outer radii of the tube \( r_i \) and \( r_o \). It can be straightforwardly calculated using the equations given in [20]. Control of wrist bending requires the inversion of (2), however this has no analytic inverse. A simple linear approximation can be derived assuming a small bending angle:

\[
\theta \approx \frac{\Delta l}{\bar{y} + r_i}
\]

(4)

Note that the spacing between the cutouts, \( x \), does not affect the bending angle but impacts the overall length of the steerable section.

C. Endoscope Fabrication

We fabricated the prototype shown in Fig. 6, which was made from a Nitinol tube with an OD of 1.8 mm and an inside diameter (ID) of 1.6 mm. This was the closest available stock size to the 1.4 mm diameter minnieScope at the time of purchase. Note that in a clinical version of this design, we would remove the outer rubber cladding from the minnieScope and replace it with Nitinol. Thus, in principle the OD of the nitinol bending section could be as small as the current OD of the minnieScope, i.e. 1.4 mm in diameter. However, custom sized nitinol tubing is expensive for initial lab prototyping, so we proceeded with the closest available stock size, and also did not remove the minnieScope cladding so as not to damage the scope accidentally during first proof of concept experiments. Also, since the endoscope was slightly larger than what we anticipate the final clinical version to be, we scaled up the middle ear phantom.
used for the validation of the device, as we shall see in the next section. The scaling factor was 1.25, which is approximately the same factor with which the endoscope was scaled due to use of stock nitinol tube sizes.

Cutouts were created in the Nitinol tube using the CNC fabrication process described in [20]. The height \( h \) and distance between cutouts \( x \) were set to 1 mm (refer to Fig. 5), while the width \( w \) was set to 1.60 mm. These parameters were determined using the mechanics model presented in [20], and in such a way to limit the maximum strain during bending below 8% and avoid plastic deformation of the device (Nitinol is typically quoted at 8–10% recoverable strain). We created a total of 8 cutouts, resulting in a steerable section of 15 mm, and an overall tip length of 17 mm (a 2 mm section was left at the beginning of the tube to accommodate the chip-tip camera). The number of cutouts was selected based on the ideal tip length derived earlier in Section II (13.45 mm), the scaling factor of the prototype \((1.25 \times)\), and the closest approximation to that value that could be created with \( x + h = 2 \) mm increments.

A 0.22 mm Nitinol tendon was used for wrist actuation. While in principle this tendon could be welded in place, to simplify fabrication it was looped around the distal cutout so that two lengths of wire run through the tube to its base. Given the tolerance between the camera and Nitinol outer sheath, the wire thickness creates a tight fit, so that the camera is held in place with a simple press fit.

We calibrated our device by applying tendon displacements in increments of \( \Delta l = 1 \) mm and measuring the corresponding bending angle. Results are shown in Fig. 7. A linear regression was found to model the behavior of the wrist between 0° and 160°, with a root-mean-square error (RMSE) of 3.71°. Our prototype thereby satisfies and exceeds the bending requirements previously derived in simulation (74.8°).

IV. PHANTOM EXPERIMENT

The endoscope prototype was tested using a custom manufactured ME phantom. The phantom was made in plastic via stereolithography 3D printing (Proto Labs, Inc., Maple Plain, MN, USA; material: Accura 60, high resolution build). The phantom shape was based on a 3D mesh reproducing real human anatomy and was generated from Computed Tomography (CT) images [2], [22], as illustrated in Fig. 8. The size was scaled up \( 1.25 \times \) to compensate for the increased size of the endoscope prototype. The phantom was made with fiducial points (see reference divots in Fig. 8) that enable the location of the robot to be registered to both the phantom and the CT scan space.

We conducted an experiment that consisted of deploying the robot into the ME phantom via the Eustachian tube and then maneuvering it via both insertion and bending of the tip. The robot was registered to the fiducial markers using camera images (as those shown in Fig. 9), which measured both the position and orientation of the endoscope tip. These photographs are then registered to the CT scan from which the phantom was generated and used to calculate the surface area viewable by the robot. During insertion, the robot was steered towards the sinus tympani (ST), to assess the attainability of the visualization rates achieved in simulation (Table I).

A Computer Numeric Control (CNC) machine was used to align the endoscope with the phantom and the endoscope was positioned initially at the eustachian tube entrance to the ME. The endoscope was then advanced into the ME phantom towards the ST using a single axis of the CNC machine and the steerable tip was curved by pulling the actuation tendon using a manually-operated positioning slider (Unislide, Velmex Inc., Bloomfield, NY) with 0.025 mm resolution. A video of the insertion was recorded and several frames were extracted for analysis (Fig. 9). The centroids of the fiducial points in the video image were
The visualization coverage achieved in the phantom experiment represents a substantial increase over what is possible with current instrumentation. In a recent study, Zhang et al. have determined that only 6.9% of the ST would be visible with a straight scope positioned at the entrance of the ET [2]. The availability of tip articulation enables the exploration of difficult-to-reach cavities like the ST, facilitating the surveillance of cholesteatoma and other ME diseases. The results presented here demonstrate for the first time the feasibility of ME endoscopy through the ET.

The results of the kinematic requirements study presented in Section II indicate that the maximum visual coverage that can be achieved in the ST varies by patient and depends on the specific ST anatomy. Further qualitative review of the simulation results along with the CT scans revealed that two important factors are the depth and morphology of the ST cavity. Most cavities can be simply inspected with a 1-DoF steerable design, especially if the endoscope can achieve a good visual alignment with the bottom of the ST, as in the case of Fig. 9. On the other hand, some patients might present a particularly deep ST, or a complex morphology that makes visualization difficult. A recent study by Marchioni et al. [30] reviewed the anatomical variations of the ST in a total of 296 ears, and found that only a limited number of cases (4.4%) present deep/complex ST shapes. These results suggest that the design presented in this letter will be suitable for use in the vast majority of the patient population. Further work will be required to identify an appropriate endoscope design for those patients who present a deep/complex ST anatomy. One option we will explore is the use of additional DoFs, i.e. the creation of additional compliant sections in the body of the robot, each presenting a different orientation and separate actuation. In principle, this would permit the creation of more complex curvilinear shapes, thereby enhancing the overall dexterity of the device. We hypothesize that having an additional DoF at the robot’s tip will enable a better alignment between the device and the ST orientation, resulting in a more comprehensive visual coverage.

Future work will focus on the automatic control of the endoscope movements and the development of a human-robot interface. It is also worth noting that this device might be possible to deploy as a non-robotic hand-held tool, as well. We envision the creation of a robotic system that incorporates image guidance and position tracking to implement automatic exploration of the ME internal volume. Such a system will integrate preoperative medical imaging (CT scans), and automatically conduct an exploration of the ME, focusing on the region of interests indicated by the surgeon in a preoperative plan. Our hypothesis is that this approach will facilitate comprehensive visual coverage of the ME, as opposed to simple manual teleoperation.

To validate the proposed device and procedure in a more realistic clinical setting, we plan to conduct cadaver studies. One important aspect that will have to be investigated is the quality of the endoscopic images provided by the chip-tip camera. The camera included in the current version of our endoscope offers a resolution of 160,000 pixels [18], which exceeds the resolution...
of prior endoscopes that have been used for middle ear procedures, e.g. the 1.6 mm endoscope used in [17] which offered 6,000 pixels. It seems intuitive that higher resolution will enable better image quality, although further studies will be needed to assess the diagnostic value of the images returned in a clinical study.

VI. CONCLUSION

The miniature robotic endoscope proposed in this letter is small enough to be passed through the eustachian tube, and its steerable tip enables it to achieve a significantly higher visual coverage of the middle ear than what would be achievable with current instrumentation. In the future, this device is expected to significantly simplify the diagnosis and treatment of patients with ME disease.

The same device can in principle be used in other medical applications where the target anatomy cannot be reached by a simple straight path, and the size of the endoscope must be as small as possible due to anatomic constraints. Potential applications include constrained anatomical environments like the brain, biliary ducts, and urinary system.

Future work will be directed towards clinical validation of the proposed device. This will require the creation of a control interface for the physician, as well as cadaver studies to further assess the diagnostic capabilities of the device. Furthermore, we will use computational simulation to further explore the design parameter space. The dexterity of the device can be further enhanced through additional DoFs, i.e. by creating a cutouts at various angles about the shaft centerline. This will enable the device to achieve more complex shapes, e.g. S shapes, which may be beneficial in increasing visualization further in the middle ear and/or facilitating applications in other areas of the body.

ACKNOWLEDGMENT

The authors are grateful to Trevor L. Bruns and Dr. Philip J. Swaney, who supported the fabrication of the endoscope prototype.

REFERENCES


