Eyes in Ears: A Miniature Steerable Digital Endoscope for Trans-Nasal Diagnosis of Middle Ear Disease

Joshua Gafford[1], Michael Freeman[3], Loris Fichera[2], Jack Noble[1,4], Robert Labadie[3,4], Robert J. Webster III[1,3,4]

[1]Vanderbilt University Engineering Department, Nashville, Tennessee, USA

[2]Worcester Polytechnic Institute, Worcester, Massachusetts, USA

[3]Vanderbilt University Medical Center, Nashville, Tennessee, USA

[4]Vanderbilt Institute for Surgery and Engineering (VISE), Nashville, Tennessee, USA

abbreviated title: Trans-Nasal Endoscope for Diagnosis of Middle-Ear Disease

correspondence: Joshua Gafford, Vanderbilt University, Nashville, TN, USA

e-mail: joshua.b.gafford@vanderbilt.edu

phone: (817)-320-3328

Abstract

The aim of this work is to design, fabricate and experimentally validate a miniature steerable digital endoscope that can provide comprehensive, high-resolution imaging of the middle ear using a trans-nasal approach. The motivation for this work comes from the high incidence of middle ear diseases, and the current reliance on invasive surgery to diagnose and survey these diseases which typically consists of the eardrum being lifted surgically to directly visualize the middle ear using a trans-canal approach. To enable less-invasive diagnosis and surveillance of middle ear disease, we propose an endoscope that is small enough to pass into the middle ear through the Eustachian tube, with a steerable tip that carries a 1 Megapixel image sensor and fiber-optic illumination to provide high-resolution visualization of critical middle ear structures. The proposed endoscope would enable physicians to diagnose middle ear disease using a non-surgical trans-nasal approach instead, enabling such procedures to be performed in an office setting and greatly reducing invasiveness for the patient. In this work, the computational design of the steerable tip based on computed tomography models of real human middle ear anatomy is presented, and these results informed the fabrication of a clinical-scale steerable endoscope prototype. The prototype was used in a pilot study in three cadaveric temporal bone specimens, where high-quality middle ear visualization was achieved as determined by an unbiased cohort of otolaryngologists. This is the first paper to demonstrate cadaveric validation of a digital, steerable, clinical-scale endoscope for middle ear disease diagnosis, and the experimental results illustrate that the endoscope enables the visualization of critical middle ear structures (such as the epitympanum or sinus tympani) that were seldom or never visualized in prior published trans-Eustachian tube endoscopy feasibility studies.

keywords: minimally-invasive surgery, endoscopy, middle ear, cholesteatoma

1 Introduction

Diseases of the middle ear (ME) affect 1 in 25 Americans every year[15]. One such disease is cholesteatoma, which is an aggressive, invasive benign tumor, commonly originating in the middle ear, that slowly and irreversibly erodes critical structures responsible for the transmission of sound. Cholesteatoma afflicts more than 40,000 Americans annually[22], and if left untreated, can cause significant morbidity including hearing loss, facial paralysis, chronic infection, meningitis, and even death in rare cases[26]. Cholesteaomas are almost exclusively managed surgically via tympanomastoidectomy[33], in which the surgeon drills away the mastoid bone to expose the middle ear cavity and surgically removes the cholesteatoma in a piecemeal fashion[16], with or without reconstruction of the ear canal and eardrum. However, cholesteatomas are difficult to diagnose, since they do not typically exhibit externally visible symptoms. As such, cholesteatomas are usually missed in primary care, only to be diagnosed later on by otolaryngologists after the disease has progressed to the point of causing permanent damage to sensitive auditory structures[5]. Timely diagnosis and treatment of cholesteatoma is critical to ensuring patient quality-of-life by halting disease progression and limiting the overall morbidity of the disease.

Due to poor differentiation of cholesteatoma from surrounding soft-tissue in computed tomography (CT) imaging[2,1,13], and poor anatomical and spatial discrimination in magnetic resonance imaging (MRI)[24,31], the only way to definitively diagnose cholesteatoma is through direct visual inspection, which is achieved by cutting the eardrum (tympanoplasty) and using an endoscope or microscope to view the middle ear through the external auditory canal[32]. Due to the rate of post-surgery cholesteatoma recurrence, most patients (about 60%[6]) must undergo staged second-look procedures (using this trans-canal approach) to surgically remove residual disease or rule out recurrence[17]. However, residual cholesteatoma is discovered in only 37% of second-look surgeries[20]. This means that the majority of cholesteatoma patients who undergo invasive follow-up surgery do not actually need it, resulting in unnecessary trauma to the patients and nearly doubling the cost of care due to the surgical expenses associated with the second-look tympanoplasty[6].

To spare patients from the trauma of invasive trans-canal second-look surgeries, and to enable clinicians to offer a less-invasive and less costly follow-up surveillance option to their patients, in this work we propose a non-invasive trans-Eustachian tube approach to ME surveillance. Instead of surgically incising the eardrum to access the ME space, a trans-Eustachian tube approach would enable visualization of the ME through the nasal passage without requiring surgery at all. While the prospect of trans-Eustachian tube middle ear endoscopy has been explored in prior work[21,25,36,8,19], feasibility studies have been carried out using only simple fiber-based endoscopes which, at only 3000-6000 pixels, lack the resolution to provide high-quality imaging of the middle ear which is critical to differentiating and diagnosing cholesteatoma. Even more importantly, the lack of a steerable tip has limited the visual field attainable by these scopes, preventing the scopes from accessing deeper regions of the middle ear where cholesteatoma is likely to occur.

While there have been several recent engineering contributions to middle- and inner-ear procedures[7], including robotic cochlear implantation[4,3], robotic mastoidectomy[9] and middle-ear microsurgery[27], there has been very little work on devices for trans-ET middle ear diagnostics and surveillance[12]. The main contributions of our work include the computational and mechanical design, fabrication, and cadaveric validation of a miniature, steerable endoscope, with high-resolution digital imaging, that can be used to diagnose middle ear disease using a trans-Eustachian tube (ET) approach. The device that we have developed, shown in Fig. 1, solves the fundamental limitations of prior attempts at trans-ET endoscopy by (1) featuring state-of-the-art complementary metal-oxide semiconductor (CMOS) camera technology to enable high-resolution (1 Megapixel) imaging of critical ME structures, and (2) implementing distal steerability to enable the deflection of the camera into regions of the ME that cannot be accessed with straight scopes[19], such as the epitympanum and sinus tympani, where cholesteatomas are most likely to occur[14]. The work described in this paper extends our early *in vitro* feasibility studies[12], with the aim of demonstrating the system concept through the development of a fully-functioning prototype and experimental validation in real cadaver tissue. This is the first manuscript to describe the cadaveric validation of a clinical-

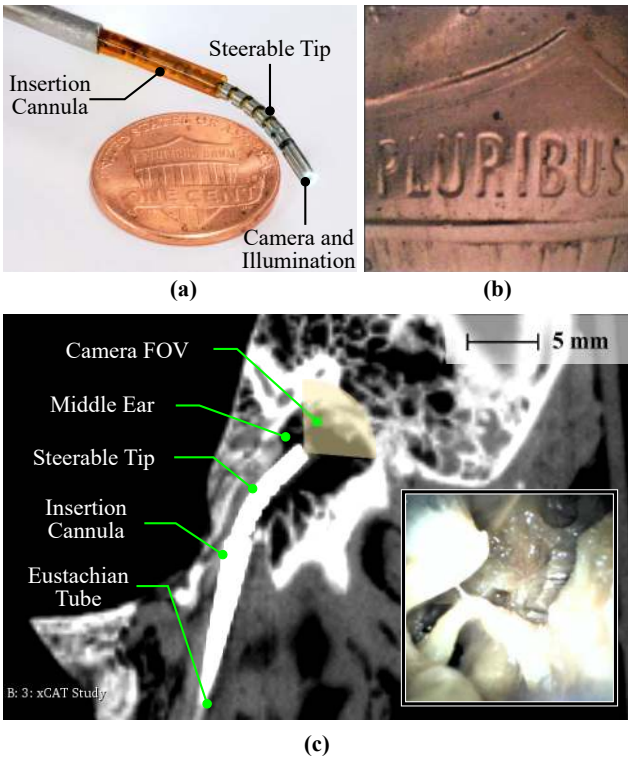

Figure 1: Miniature steerable endoscope for middle-ear surveillance: **(a)** close-up of the distal steerable section extending out of the insertion cannula, which features a chip-on-tip camera and fiber illumination integrated into a micromachined Nitinol tube (a US penny provides scale), **(b)** an image of a penny detail as captured by the camera, **(c)** CT scan of the steerable endoscope surveying the middle ear of a temporal bone specimen, where the inset shows the view as seen through the endoscope.

scale endoscope, with digital imaging and a steerable tip, designed to pass through the ET, and is a fundamental step towards office-based non-surgical diagnosis of cholesteatoma and other ME diseases.

2 Materials and Methods

In this section, a computational framework for the design of a steerable endoscope for middle ear surveillance is presented. This framework consists of a simple kinematic model of the steerable tip, coupled with a sampling-based motion planning and ray-casting algorithm, and was employed to assess the ability of the steerable scope to visualize the middle ear from segmented CT scans of real human patients. The results of this framework demonstrate

that a steerable scope provides superior visual coverage over rigid/non-steerable scopes, and these results informed the design and fabrication of a functional prototype of the steerable endoscope (shown in Fig. 1(a)). The endoscope prototype was validated through a series of experiments in cadaveric temporal bones specimens which demonstrate the scope's ability to visualize critical ME structures using a trans-ET approach (shown in Fig. 1(c).)

2.1 Design and Modeling of the Endoscope's Steerable Tip

The deflection of the steerable tip in our endoscope is enabled by the use of a distal tendon-actuated flexural bending mechanism (Fig. 2). This design was first proposed by Swaney and York et al.[34,37], and is realized by a tube of superelastic Nickel-Titanium alloy (Nitinol) with a series of laser-cut asymmetric rectangular notches machined into the tip. Removing material from the tube in this manner offsets the neutral bending axis of the tube (Fig. 2 (c)), resulting in bending of the structure when actuated by a single pull-wire (Fig. 2 (b)). Due to Nitinol's superelastic properties, when the tension on the pull-wire is removed, the tube returns to its original undistorted configuration. As we demonstrate in the following section, it is possible to establish a simple mathematical mapping between the amount of displacement applied to the pull-wire and the resulting tube pose[34,37]. This notched joint design is particularly suitable for our application for the following reasons: (1) bending requires a single pull-wire, leaving most of the internal lumen of the tube available for the passage of the chip-tip camera, optical fibers and electrical wires; (2) the joint can be easily made in tubes of diameter < 2 mm (as it was demonstrated in prior work[37,34,12]) — this is required to enable passage of our instrument through the tiny channel of the Eustachian tube; (3) the mechanism has been shown to be robust to thousands of bending cycles[34]; and (4) the offset neutral axis creates the longest possible moment arm between the flexural backbone and the location where the pull-wire applies force, enabling much lower tendon tension forces to create bending when compared to more traditional designs where the neutral axis is in the center of the shaft. In the following, we briefly review the kinematics of this bending mechanism (the reader is referred to prior work by York et al.[37] and Swaney et al.[34] for a

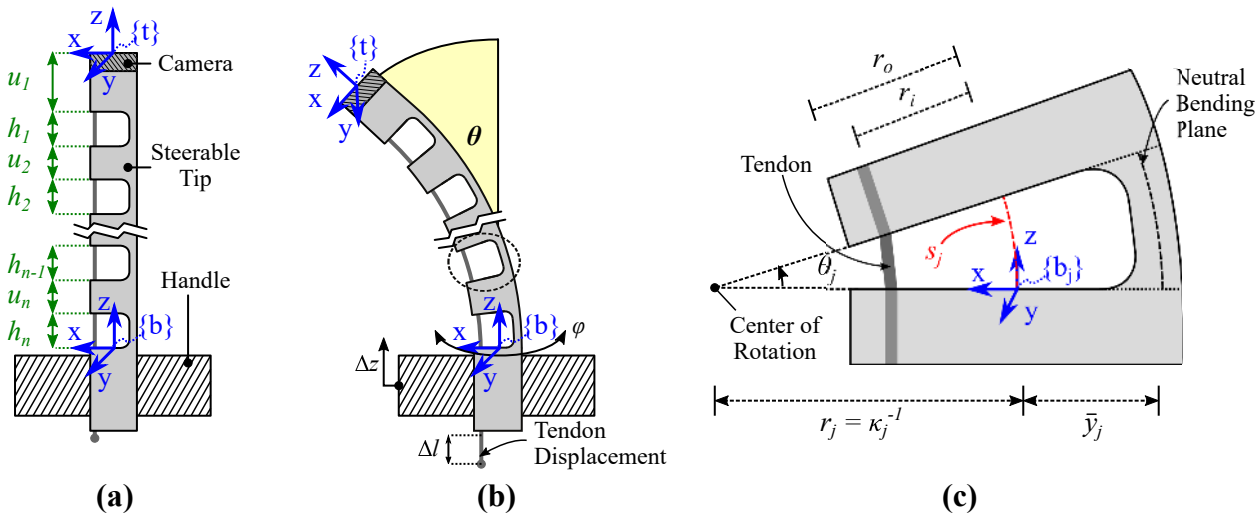

Figure 2: Kinematics of the notched steerable tip: **(a)** asymmetric notched tube and relevant geometric parameters (where u_j is the j^{th} solid section length, h_j is the j^{th} notched section length, and x, y and z denote the local coordinate axes), **(b)** illustration of steerable tip deflection θ when the pull-wire is displaced by Δl, with additional actuation parameters denoted (where Δz is the axial translation and ϕ is the axial rotation angle of the endoscope), and **(c)** close-up of the j^{th} notch with relevant geometric parameters (where r_o and r_i are the tube outer and inner radii (respectively), r_j is the bending radius, s_j is the arc length, θ_j is the bending angle, κ_j is the curvature, and $\bar{y}_j$ is the neutral axis location of the j^{th} section.)

full derivation of the kinematics). This kinematic model is used in a simulation study which shows that distal steerability enables comprehensive visualization of parts of the middle ear that would otherwise be impossible to visualize with a rigid, non-steerable scope.

2.2 Kinematic model

The kinematics of the notched steerable section can be calculated by modeling the Nitinol tube as a sequence of flexible (notched) sections and solid (uncut) tube sections with length h_j and u_j respectively, where $j \in \{1, ..., n\}$ and n is the total number of notches (refer to Fig. 2 (a)). For this analysis, SI units are assumed, and we prescribe local and global coordinate frames as defined in Fig. 2 (where z points along the axis of the tube, y is parallel to the axis of bending, and x completes the frame according to the right-hand convention). Let us first consider the j^{th} solid (uncut) section. Assuming rigidity (i.e. no flexural deformation),

the homogeneous transformation matrix $T_{s,j}(u_j)$ associated with this section consists of a simple translation of length u_j along the z axis, i.e.,

$$T_{s,j}(u_j) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & u_j \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{1}$$

The kinematics for the j^{th} notched section, which undergoes flexural bending, depend on the amount of tendon displacement Δl (refer to Fig. 2(b)): intuitively, the more the tendon is pulled, the tighter the nothed sections will close. Following the approach described in[37,34], we assume that notched sections deform in the shape of a constant curvature arc, and map the amount of tendon displacement Δl to the arc parameters κ_j and s_j with the following relations:

$$\kappa_j \approx \frac{\Delta l}{h_j(r_i + \bar{y}) - \Delta l \bar{y}} \tag{2}$$

$$s_j = \frac{h_j}{1 + \bar{y}\kappa_j} \tag{3}$$

where h_j is the height of the j^{th} notch, r_i is the inner radius of the tube, and $\bar{y}$ is the neutral bending plane of the notch, which can be calculated using the equations in[34,37]. To obtain an homogeneous transformation for the j^{th} notch, we plug the arc parameters κ_j and s_j into the following matrix, which describes motion along a constant curvature arc in the $x - z$ plane:

$$T_{f,j}(s_j, \kappa_j) = \begin{bmatrix} \cos(\kappa s) & 0 & \sin(\kappa s) & \frac{1-\cos(ks)}{k} \\ 0 & 1 & 0 & 0 \\ -\sin(\kappa s) & 0 & \cos(\kappa s) & \frac{\sin(\kappa s)}{k} \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4}$$

Finally, the transformation from the base of the tube to its tip is given by:

$$T_{\text{b}}^{\text{t}} = \prod_{j=1}^{n} (T_{s,j}\ T_{f,j}) \tag{5}$$

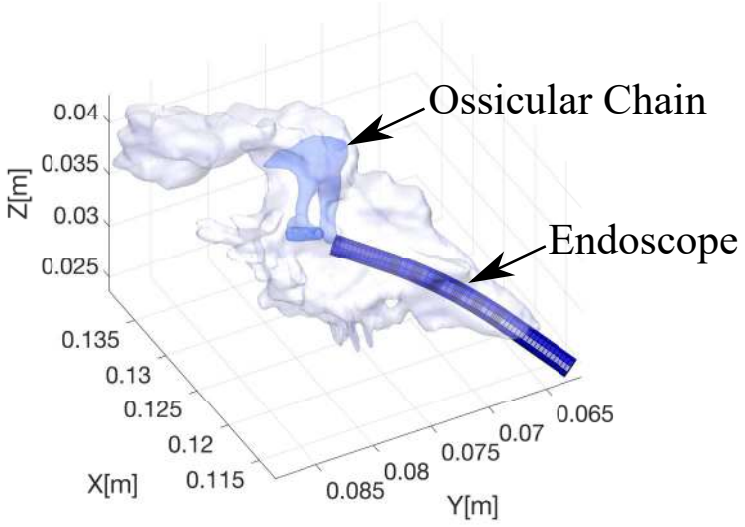

Figure 3: Simulated trans-Eustachian endoscope deployment into the middle ear.

The net result is a model that, for a given initial base pose T_{b} and proximal displacement of the tendon by some distance Δl, calculates the position and orientation of the camera at the tip of the steerable section. Note that the model above does not account for axial rotation ϕ or translation Δz of the endoscope, but these motions can be easily incorporated by pre-multiplying T_b^t with the relative basic transformation matrices.

2.3 Computational Simulation Studies

Using the simulation environment created in prior work[12], we simulated the ability of a 6-notch tube to visualize the middle ear when emerging from the Eustachian tube, as shown in Fig. 3, based on the kinematic model from the previous section. We performed simulations on a total of $n = 6$ geometric middle ear models generated from CT scans of real patients. These models were manually segmented by a physician to identify different sub-regions of the middle ear cavity (Fig. 4), using the methods described by Zhang et al[38].

Endoscope motion was simulated through a sampling-based motion planning algorithm (rapidly-expanding random trees, or RRT[23]) to define a set of collision-free paths within the geometric middle ear model. We parameterize the endoscope configuration space as a vector

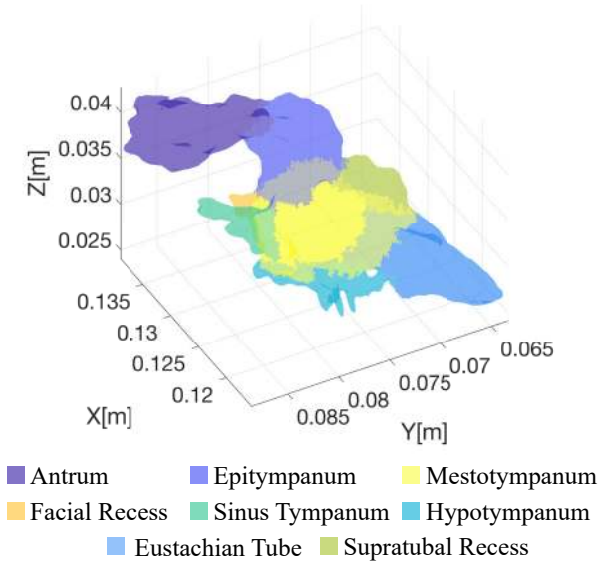

Figure 4: Computer model of the middle ear showing different sub-regions.

$\mathbf{q} \in \mathcal{Q} \subset \mathbb{R}^3$ containing φ, Δz, and Δl, where φ is the axial rotation of the endoscope (about the z axis of its base frame), Δz is the axial translation along the z-axis of the base frame, and Δl is the tendon displacement which results in an overall tip transformation based on the kinematic model derived in Section 2.2. These configuration space variables are shown in Fig. 2(b). Starting from an initial configuration $\mathbf{q}_0$, where the scope is aligned with the Eustachian tube and positioned at the entrance of the ET isthmus, RRT incrementally builds a tree of configurations that the endoscope can reach within the middle ear in a collision-free path. The RRT algorithm is known to have probabilistic completeness[23], which is a useful property in our study: it implies that the longer the algorithm runs for, the higher the likelihood of discovering the entire reachable workspace of the endoscope is, which enables us to estimate the total extent of middle ear cavity that the endoscope is able to reach and visualize.

A ray-triangle intersection procedure was used to detect and discard any configurations that resulted in collisions with the segmented anatomy. Once the set of reachable configurations was generated, the total visible surface was estimated using a computational ray-casting technique [38], with an endoscope field of view (FoV) of 90 degrees. During ray-casting, a cone

Outer Diameter (d_i)	1.62 mm
Inner Diameter (d_o)	1.40 mm
Number of notches (n)	6
Notch height (h)	0.81 mm
Notch width (w)	1.32 mm
Spacing (u)	1 mm
Maximum Deflection (θ_{max})	90 deg
Minimum Curvature (κ_{min})	0.2 mm^{-1}

Table 1: Design parameters of the notched steerable endoscope.

of rays are projected from the distal end of the simulated tube with the prescribed FoV for each of the collision-free configurations, and the Möller-Trombore algorithm[28] is used to compute intersections between the rays and the triangulated computer model of the ME. These intersections are collected and assigned to each of the segmented sub-regions shown in Fig. 4 to compute an overall percentage visualization for each respective sub-region. The geometric parameters for the simulated endoscope tip design are summarized in Table 1. Note that the maximum deflection and minimum curvature were selected based on a mechanical factor-of-safety of 2. This simulation was performed on both the steerable scope design, and a rigid scope with identical optical properties, to evaluate the benefits of distal steerability in terms of overall visualization.

2.4 Mechanical Design of Steerable Wrist

The overall steerable endoscope system consists of the steerable tip sub-assembly, the user interface sub-assembly, and peripheral electronics for illumination and image capture. The steerable wrist is comprised of a Nitinol super-elastic tube (with the same outer- and inner-diameter of the simulated tube from Table 1) and an image sensor and illumination sub-assembly, as shown in Fig. 5. The outer diameter of the wrist was selected to enable passage

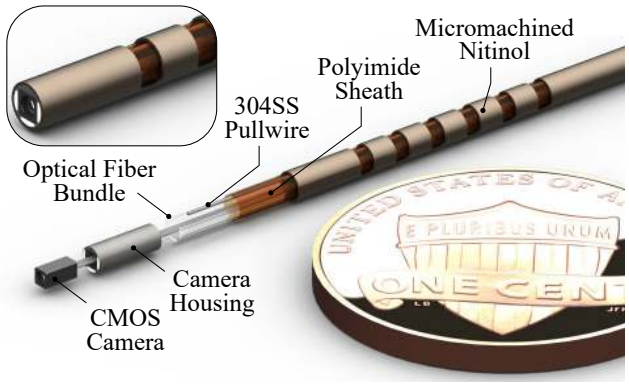

Figure 5: Exploded view of the steerable wrist sub-assembly (shown with a US penny for scale) where the inset shows the camera sub-assembly integrated into the Nitinol wrist.

into the ME through the narrow bony isthmus of the Eustachian tube which measures 1.7-1.9mm in diameter on average[30]. To create the series of distal notches, the stock Nitinol tube (P/N 84237, Johnson Matthey) was glued into a slotted sacrificial aluminum jig with cyanoacrylate, and a 0.81mm end-mill was used to machine the notches via CNC milling. The cyanoacrylate was dissolved in acetone to release the notched Nitinol tube from the sacrificial fixture, which was then lightly sanded with 1200-grit sandpaper to remove any leftover burrs. These machined notches offset the neutral axis of the tube such that a tensile force created at the distal end by a pull-wire results in smooth constant-curvature bending of the tip (as described in Section 2.1 and illustrated in Fig. 2).

After the notched Nitinol tube was fabricated, it was integrated with the image-sensor sub-assembly, which consists of a CMOS image sensor with integrated optics (minnieCam-XS, Enable Imaging Inc., Redwood City, California), three optical fiber bundles, and a 150μm 304 stainless steel pull-wire which are all terminated within a custom housing (manufactured via wire-EDM) and fastened with epoxy. The CMOS image sensor features a resolution of 1 Megapixel, a 90-degree field-of-view, 60Hz frame rate, and an overall diagonal footprint of 1.3mm. The image sensor, fiber optic bundle, and steel pullwire were epoxied into the stainless-steel housing, and the entire image sensor assembly was then press-fitted into the distal end of the Nitinol tube, with a thin polyimide tube inserted in between the optical/electrical wires and the Nitinol tube to further encapsulate and protect the wires. The

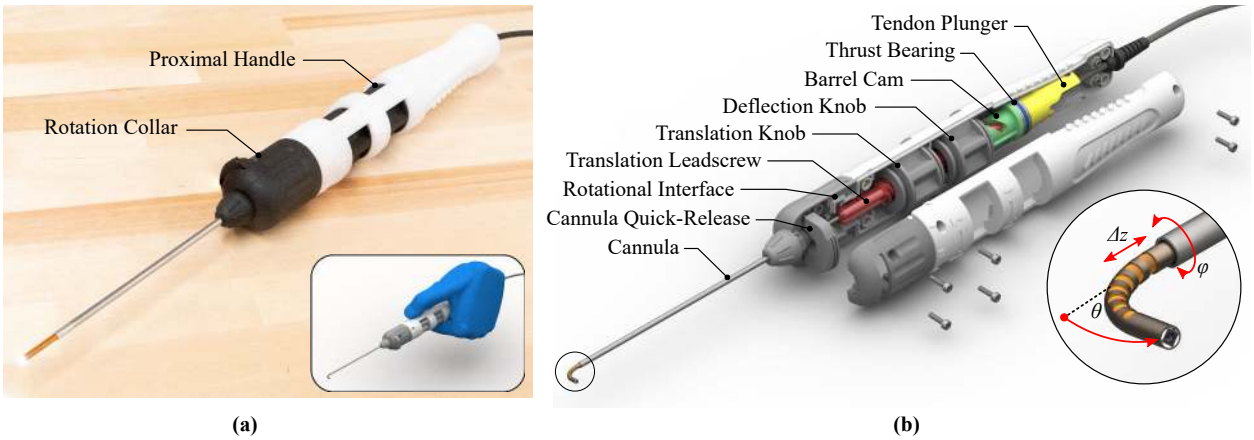

Figure 6: An ergonomic handle was created to enable physicians to control the camera in three degrees of freedom: **(a)** photograph of the system, where the inset shows how the physician will grasp the device, **(b)** Rendering of the handle design with callouts to critical features the inset shows a close-up of the steerable tip (and actuatable degrees-of-freedom which include Δz (insertion/retraction), ϕ (axial rotation), and θ (tip deflection)).

image sensor wire leads and optical fibers extended out of the proximal end of the tube and were terminated with standard optical and video connectors (SMA fiber-optic coupler and micro-USB, respectively). A Pulse-Width Modulated (PWM) LED driver (lightLume -M, Enable Imaging Inc.) provides a source of illumination that gets transferred to the distal end via the optical fiber bundle. A video processing unit (VPU-USB3-HDMI-XS, Enable Imaging Inc., with automatic gain adjustment and white balancing) converts the image sensor signal to an HDMI output for display on standard video monitors.

2.5 Mechanical Design of the User Interface

An ergonomic user interface was designed to enable the physician to control (1) the insertion and retraction of the camera into and out of the middle ear (Δz), (2) the rotational orientation of the camera inside the middle ear (ϕ), and (3) the tip deflection (θ).

The handle interface design is shown in Fig. 6. The physician holds the device using a power grip for stability (shown in the inset of Fig. 6 (a)), and the two rotary input knobs are manipulated with the thumb and forefinger via a pinch grasp for precise actuation. The

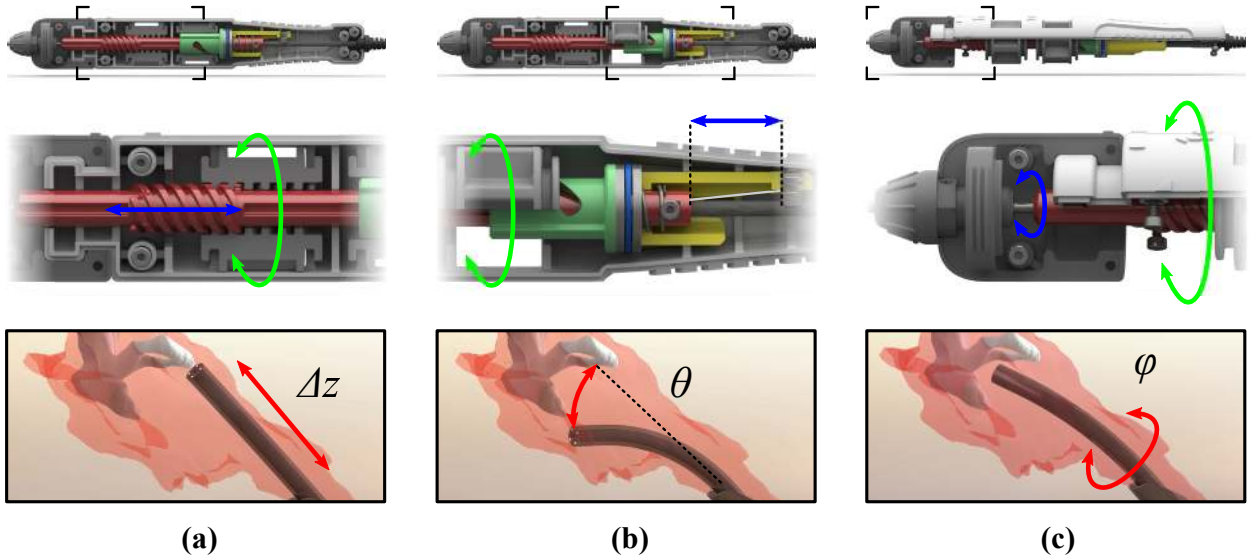

Figure 7: Functional illustration of ergonomic handle (from top: section view of the handle, detailed view of the actuated degree-of-freedom, and the resulting motion of the steerable tip within an anatomical rendering of the middle ear space): **(a)** rotation of the distal input knob translates the keyed leadscrew to actuate the tip in Δz, **(b)** rotation of the proximal input knob rotates a keyed barrel cam to apply tension to the pull-wire and actuate the tip in θ, and **(c)** rotation of the entire proximal handle with respect to the rotation collar actuates the tip in ϕ.

insertion/retraction degree-of-freedom is enabled by a non-backdriveable leadscrew which couples the first (translation) input knob to the linear translation of the steerable endoscope with respect to the insertion cannula (as shown in Fig. 7 (a)). The rotation of the leadscrew is constrained by a keyway in the leadscrew which engages with keys in the handle body, causing the leadscrew to translate linearly when the input knob is rotated. The lead p of the leadscrew is 12.7mm, enabling fine, smooth, and controlled insertion and retraction of the steerable section into and out of the middle ear space with two turns of the translation knob. Satisfying the condition of non-backdriveability ensures that the force generated from the deflection degree-of-freedom does not influence the position of the leadscrew. For leadscrew transmissions, the condition for non-backdriveability is as follows:

$$p > \frac{2\pi r_s \mu_s}{\cos(\alpha_p)} \tag{6}$$

where r_s is the screw radius, μ_s is the friction between the leadscrew and leadnut (in this

case, the translation knob), and α_p is the thread angle (in this case we have implemented a power screw design with α_p=90°).

The overall deflection of the steerable tip (θ) was designed to be decoupled from the insertion and retraction degree-of-freedom. As the leadscrew translates distally by turning the first input knob, the slotted barrel cam is drawn into the deflection knob via keyway engagement. As the second (deflection) input knob is rotated, the barrel cam rotates about the axis of the leadscrew (due to the keyway constraint which couples the deflection knob rotation to the barrel cam rotation) and translates axially along the leadscrew (due to a diagonal slot-pin cam interface which pushes the barrel cam backwards as the deflection knob is rotated). As the barrel cam is driven backwards due to the rotation of the deflection knob, a thrust bearing accepts the axial load created by this translation, while permitting rotation of the barrel cam as it compresses a spring-loaded plunger mechanism which applies tension to the pull-wire with respect to the leadscrew. This is shown in Fig. 7(b). The pull-wire terminates at the distal end of the scope, such that tension in the pull-wire causes the steerable tip to deflect, as illustrated conceptually in Fig. 2(b). When the input torque is removed from the deflection knob, a bias spring returns the mechanism to its neutral (undeflected) state, and the superelasticity of the Nitinol steerable tip enables the camera to return to its undeflected configuration. As the barrel cam mechanism can be driven only by rotational motion with respect to axis of the leadscrew, the purely axial translation of the leadscrew does not couple to the deflection degree of freedom.

The deflection of the tip θ is related to the rotation of the input knob θ_{rot} as shown in Equation (7):

$$\theta = \frac{r_{cam}\theta_{rot}\tan(\beta)}{\bar{y} + r_i} \tag{7}$$

where r_{cam} is the radius of the barrel cam and β is the angle of the cam slot with respect to the leadscrew axis. Similarly, we can determine the input torque required to deflect the tip,

as follows:

$$\tau_{rot} = \frac{r_{cam}(k_{tip} + k_{bias})\theta(\bar{y} + r_i)}{\tan(\beta) - \mu_c \sin(\beta)\cos(\beta)} \tag{8}$$

where k_{tip} and k_{bias} are the bias forces of the steerable tip and the return spring (respectively), and μ_c is the friction coefficient of the cam interface. The cam angle β determines the motion and torque scaling between the input rotation and the output deflection, which are inversely correlated and should be selected so that the tip deflection can be comfortably actuated by a rolling motion between the thumb and fore-finger. In this prototype, a cam angle of $\beta = 45°$was chosen, resulting in a maximum knob rotation of 25°, and an input torque of 25 N·mm. These are well within recommended ergonomic limits for rotational controls actuated by the thumb and forefinger[10].

The camera's rotation is enabled by a simple journal bearing interface between the proximal handle component (the portion that houses the input knobs) and a distal rotation collar (the portion that houses the cannula quick-release), as shown in Fig. 7(c). An insertion cannula is kinematically decoupled from the steerable tip, allowing the tip to be withdrawn inside the insertion cannula to protect the camera while positioning the scope into the Eustachian tube. A spring-loaded ball detent mechanism creates a tactile and audible 'click' when the handle is rotated into its home position, and a series of fiducial markers embossed into the handle components enable the physician to observe the extent of the rotation. By observing the fiducial configuration, the physician can mentally register the endoscope view to the patient's anatomy. The rotation mechanism enables the physician to keep the insertion cannula stationary inside the patient's nasal passage while rotating the steerable camera within it. A spring-loaded cannula quick-release mechanism enables the physician to quickly swap insertion cannulas of different diameters and angles on an as-needed basis.

2.6 Pilot Validation in Temporal Bone Specimens

To demonstrate the efficacy of our ME scope in a clinically realistic model, we performed a series of visualization experiments in cadaveric temporal bone specimens. Three formalin-

preserved human temporal bone specimens, transected through the nasal cavity to preserve the entrance to the Eustachian tube, were acquired for the preliminary validation of the trans-ET endoscope proposed in this paper. Prior to each of the surveillance experiments, the temporal bone specimens were loosely positioned into a temporal bone holder and clamped into place. A Eustachian tube balloon dilation device (Relieva SpinPlus, Acclarent) was inserted into the Eustachian tube, and the balloon was pressurized for approximately 30 seconds to dilate the ET and facilitate insertion of the endoscope cannula and minimize contact between the insertion cannula and the side walls of the ET. Once the ET was sufficiently dilated, the steerable endoscope (with the camera initially retracted back into the insertion cannula) was inserted into the Eustachian tube and advanced until the tip of the cannula was positioned just proximal to the ET isthmus. Under the supervision of an experienced otolaryngologist, the camera was then advanced in z out of the insertion cannula and into the middle ear space. For each specimen, several minutes of video were captured while the various degrees of freedom (z, ϕ and θ) were actuated within the middle ear to capture various anatomical features. Video was recorded for several minutes until the participating otolaryngologist was satisfied with the degree of visualization achieved (mean recording time: 214$\pm$30.7 seconds), and the endoscope was retracted. Computed tomography (CT) images of each temporal bone specimen were then captured using a compact, portable CT scanner (xCAT XL, Xoran Technologies, Ann Arbor, Michigan) intended for real-time scanning of the brain, sinuses, ears, and skull-base, with a minimum slice thickness of 0.1mm. A combination of automatic (using custom segmentation software developed by Noble et al.[29]) and manual (3DSlicer[11]) techniques were used to segment the relevant ME anatomy for post-processing and analysis.

Following the conclusion of the pilot experiments, the captured videos and an associated five-point Likert survey were distributed to $n = 7$ unbiased otolaryngologists (none of whom had any association with the development of the endoscope or the writing of this manuscript) with variable experience levels ranging from residents to attending physicians, all recruited from the Vanderbilt University Medical Center Department of Otolaryngology according to

approved IRB protocol 200449. Each participant was asked to watch each of the videos, and rate the overall quality of visualization by identifying several critical structures in each of the middle ear specimens where cholesteatoma is likely to occur, including the epitympanic recess, mesotympanum, ossicular chain, sinus tympani and the tympanic membrane[14]. Results from the Likert surveys that were distributed to practicing otolaryngologists were analyzed in a statistical programming software package (R), wherein a Wilcoxon signed rank test was employed to determine the statistical significance of the categorical survey responses.

3 Results

3.1 Computational Simulation Results

Using the kinematic model, RRT algorithm and computational raycasting as previously described, we estimated the potential visual coverage that our steerable scope could provide on segmented middle ears from $n = 6$ real patients. In addition, to understand the utility of having a steering tip in our endoscope, we repeated the same simulations with a rigid endoscope (i.e. an endoscope with identical optical properties that lacks distal steering). The results of this analysis are summarized in Fig. 8. The visualization performance is reported as a percentage of the overall segmented feature visualized (where each feature corresponds to the colored regions in Fig. 4).

As can be observed from Fig. 8, the integration of distal steerability significantly improves the visualization of all tested features in the middle ear for all specimens, with notable improvements in visualization of the epitympanum ($p << 0.001$), mesotympanum ($p << 0.001$), and sinus tympani ($p << 0.001$), where cholesteatoma is likely to originate. These simulations also show that it is possible to partially visualize the hypotympanum with a steerable scope, which would be impossible to visualize with a rigid scope that only offers line-of-sight visualization from the ET isthmus. Overall, these simulations corroborate our hypothesis that steerability is essential for comprehensive visual inspections of the middle ear using a trans-ET approach.

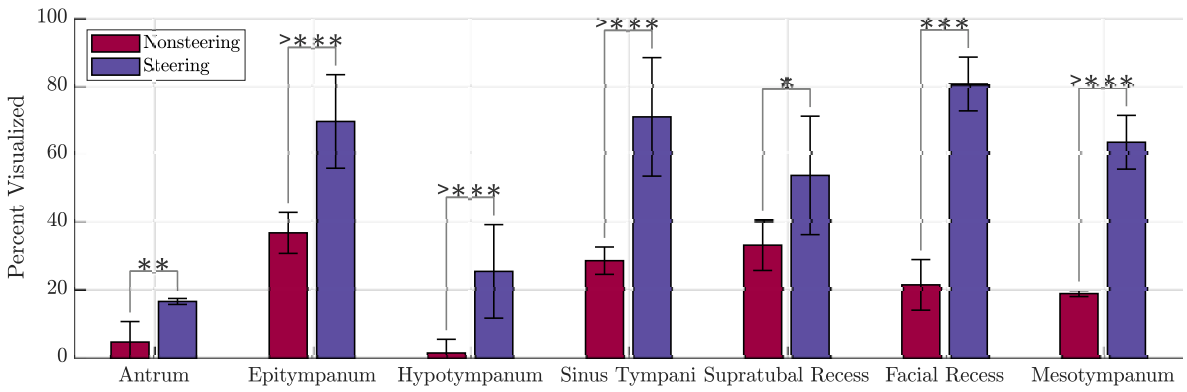

Figure 8: Simulation results comparing visual coverage of the ME offered by a rigid (non-steering) endoscope, and the endoscope that we propose in this paper. Error bars correspond to $\pm 1\sigma$ for $n = 6$ specimens. A two-sample t-test reveals that integrated steerability results in significant visualization improvements over non-steerable designs. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$

3.2 Pilot Validation in Temporal Bone Specimens

Image snapshots from the videos recorded during the temporal bone studies are shown in Fig. 9, along with the segmented CT scan of the middle ear for each of the specimens. As the images elucidate, the high resolution of the CMOS camera provides a clear view of the ME space, with many critical features immediately visible with high clarity and illumination. The results from this analysis for each of the temporal bone specimens visualized are summarized in Fig. 10. As the results demonstrate, the participants responded favorably to the middle ear surveillance videos captured via the steerable trans-ET endoscope, as indicated by the proportion of responses that are ranked 'Acceptable', 'Good', and 'Excellent.' A Wilcoxon signed rank test with Bonferroni correction was employed to test the hypothesis that the middle ear structures can be adequately visualized via the trans-ET scope (by computing the mean shift between the neutral or 'Acceptable' response and the actual distribution of observed responses), and corroborates this observation by revealing that the visualization of all of the structures can be classified as 'Acceptable' to 'Excellent' with high significance (epitympanum: $p < 0.001$, mesotympanum: $p < 0.001$, ossicular chain: $p < 0.001$, sinus tympani: $p < 0.01$, tympanic membrane: $p < 0.01$, overall visualization quality: $p < 0.001$). Of the structures surveyed, the sinus tympani proved the most difficult to visualize

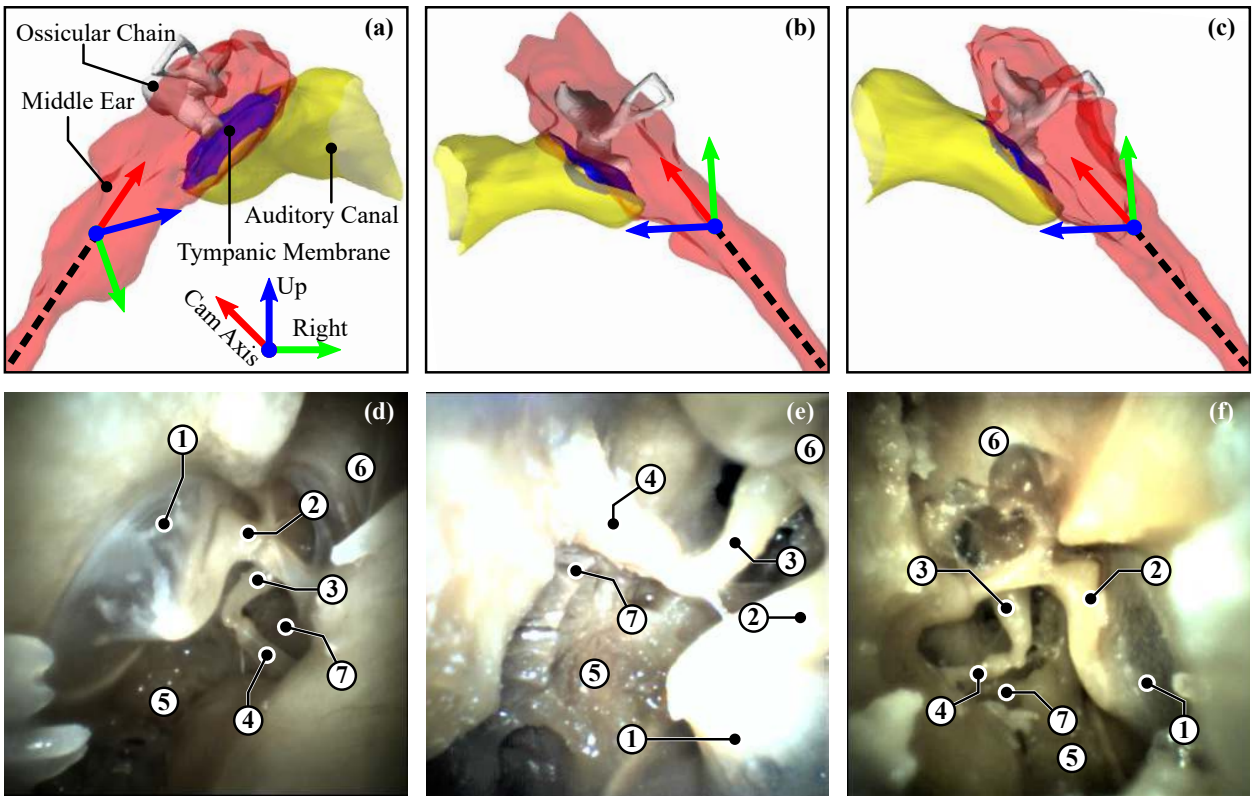

Figure 9: Trans-ET surveillance of temporal bone specimens, where **(a)-(c)** shows the segmented ME anatomy with an illustration of the camera's initial orientation, and **(d)-(f)** shows the view through the endoscope with callouts to relevant ME anatomy: ① Tympanic Membrane, ② Malleus, ③ Incus, ④ Stapes, ⑤ Mesotympanum, ⑥ Epitympanum, ⑦ Sinus Tympani.

especially in Specimen 1 (with one responder unable to visualize the sinus tympani at all). However, as the survey results and our subsequent statistical analyses demonstrate, other critical middle ear structures (ossicular chain, tympanic membrane, mesotympanum and epitympanic recess) were adequately visualized in all specimens tested.

4 Discussion

In patients with middle ear pathology, surgical exploration is often necessary to confirm diagnosis of disease, or survey for recurrent disease. The current standard-of-care requires surgical incising of the tympanic membrane to gain access to the ME space, a procedure that is both traumatic to the patient and costly for the hospital. In this work, we propose a

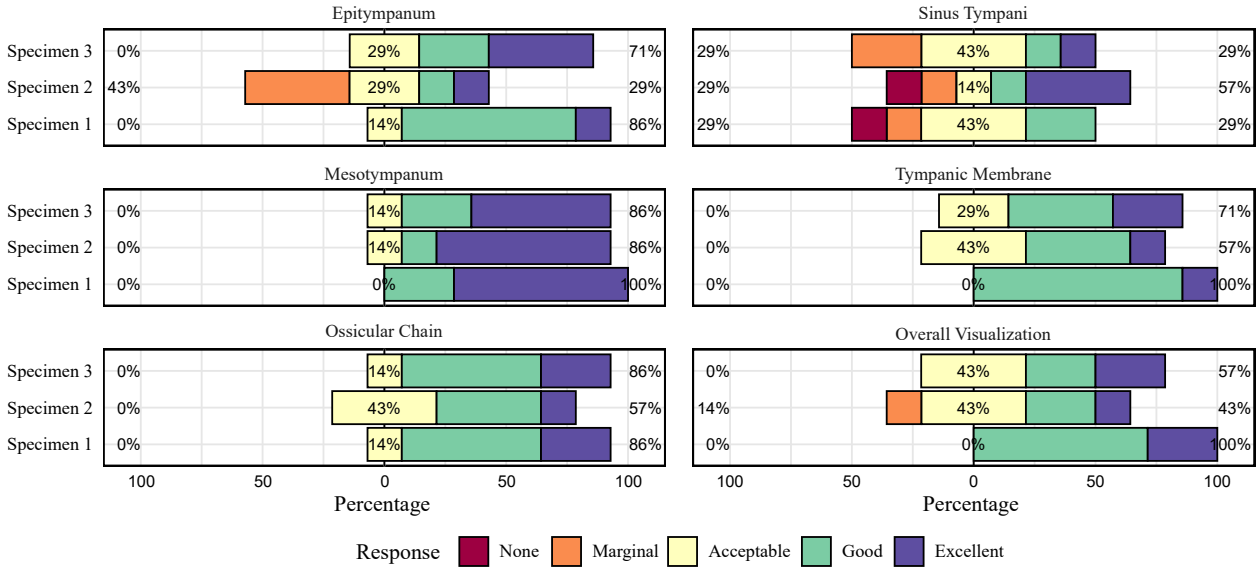

Figure 10: Likert survey results from the temporal bone experiments, grouped by structure. A Wilcoxon signed rank test reveals that the structures could be visualized via the trans-ET steerable scope with 'Acceptable' to 'Excellent' quality. Of the structures surveyed, the sinus tympani was the most difficult to visualize.

novel, steerable digital endoscope that is small enough to enter the middle ear space through the Eustachian tube, potentially sparing patients from invasive second-look surgery in favor of non-invasive, office-based procedures. This work improves upon prior work in fiber-based ME endoscopy[21,25,36,8,19] to enable high-resolution digital imaging of the ME space, with a steerable tip to enable access to anatomical features that would otherwise be impossible to visualize with rigid scopes.

Our scope has a steerable tip which affords the physician with the ability to observe structures that would otherwise be poorly visible or not visible at all with rigid non-steerable scopes. This was demonstrated computationally as shown in Fig. 8. The simulation results illustrate that the addition of distal steerability comprehensively improves the visualization of all middle ear structures tested, even enabling partial visualization of sub-regions of the ear which are impossible to visualize with a non-steering tip (such as the hypotympanum).

Through our pilot study on temporal bone specimens, we showed that the steerability afforded by our endoscope, combined with high-resolution digital imaging and fiber-based illumination, enables the surveillance of critical middle ear structures where cholesteatoma

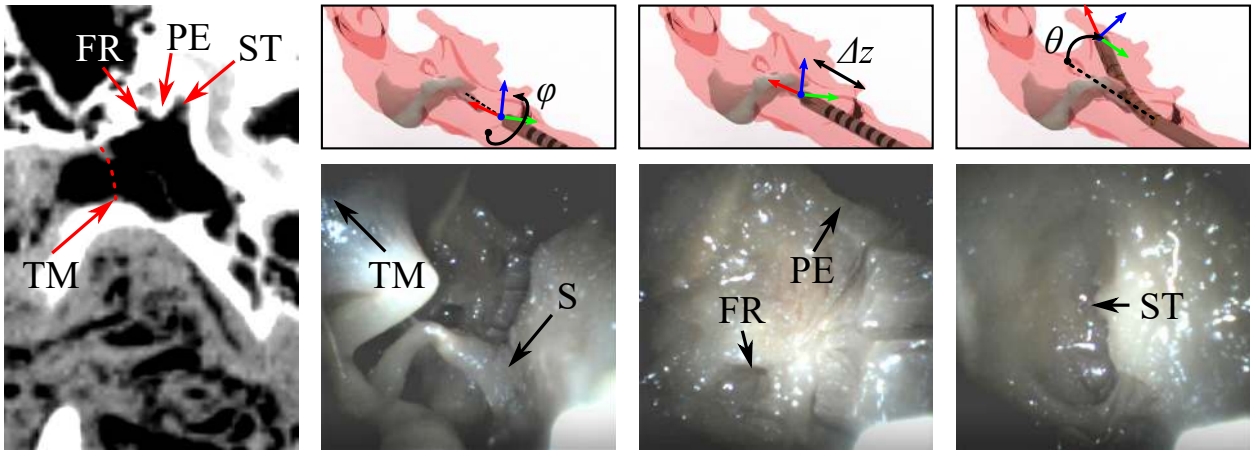

Figure 11: Visualization of the sinus tympani as enabled by the steerable tip: **(a)** CT scan of the middle ear space with callouts to relevant structures, **(b)** actuation in ϕ rolls the scope to align the field-of-view (where top shows a rendering of the scope in the middle ear, and bottom shows the endoscope image), **(c)** actuation in z inserts the endoscope further into the ME, past the tympanic membrane and ossicular chain towards the pyramidal eminence, and **(d)** deflection of the tip in θ allows for visualization of the sinus tympani.

is likely to occur. The benefit of integrated steerability is illustrated in Fig. 11, which shows a sequence of extracted video stills and a representative computer rendering of the scope's motion within the middle ear during the surveillance sequence. From the ET orifice, the scope is rotated such that the canal and inner-ear surfaces of the ME are aligned laterally. The scope is then advanced in z further into the ME, past the stapes footplate and towards the pyramidal eminence (PE) that separates the facial recess (FR) and the sinus tympani (ST). While this would be the limit of visualization for a non-steering scope, the extra steerability affords further deflection in θ around the PE to enable high-quality visualization of the ST (Fig. 11(d)) which is not in the direct line-of-sight from the vantage point of the ET isthmus (Fig. 11(b)). This is a substantial improvement over previous work in trans-ET surveillance, where attempts were made to visualize the middle ear via low-resolution fiberoscopy. Due to fundamental limitations in resolution, illumination, depth-of-field and distal dexterity, the tympanic membrane was visualized less than half of the time (44%), the sinus tympani was seldom visualized (22% of the time) and the epitympanum was never visualized[19]. From the survey responses generated from our temporal bone experiments,

the sinus tympani was visualized in 90% of responses for the three specimens, and the epitympanum and tympanic membrane were both visualized in 100% of responses. This indicates a substantial improvement over prior attempts to visualize the ME endoscopically using a trans-Eustachian tube approach.

The work has several limitations worthy of mention. The temporal bone specimens were de-identified, preventing us from observing any age or gender-specific biases. While temporal bone specimens accurately reproduce ET and ME anatomy to prove feasibility of the concept, a follow-up study in full cadaver heads is necessary to comprehensively model the full clinical workflow which includes trans-nasal dilation of the ET and subsequent passage of the endoscope into the ET. As with any endoscopic technology, debris occlusion and lens fouling is an inherent limitation. In this work, there were instances where the endoscope had to be retracted to clean the lens. However, pre-dilating the Eustachian tube helped to minimize contact between the endoscope and the ET wall which mitigated fouling and occlusion. Finally, we emphasize that the steerable endoscope as described is still in the proof-of-concept phase, and development is currently underway to address issues regarding sterilizability, designing for manufacturability and assembly, and processes for obtaining regulatory approval. Towards these aims, future work will focus on modifying the design of the steerable tip and the handle towards clinical readiness. To anticipate manufacturing scale-up, the notched steerable section will be laser-manufactured using femtosecond laser technology, passivated and electropolished under an ISO:13485-accredited quality systems framework. The handle components in the current prototype are 3D-printed from an acrylonitrile butadiene styrene (ABS) material. Following a rigorous design for manufacturing and assembly (DFMA) analysis to reduce the number of constituent components, the handle will be re-designed to minimize the number of heterogeneous parts and assembly steps, and injection-molded from a biocompatible thermoplastic such as polycarbonate to improve robustness and reduce the overall cost to manufacture at volume.

While trans-ET middle ear surveillance is an excellent initial application, there are other clinical applications that could benefit from the proposed steerable endoscope technology.

An obvious extension is endoscopic transsphenoidal resection of pituitary tumors, which uses a similar trans-nasal approach to access and visualize tumors occurring in the skull-base. Improved visualization of the surgical field during such procedures could potentially increase therapeutic yields[35]. Joint arthroscopy is another example where a miniature, needle-sized steerable scope could potentially lower learning curves and improve diagnosis and treatment, while minimizing many procedure-related complications that arise from the use of rigid/non-steerable scopes (such as iatrogenic damage to articular cartilage[18]).

5 Conclusion

This study is the first of its kind to demonstrate the efficacy of trans-Eustachian tube middle ear endoscopy using a steerable endoscope with high-resolution digital imaging. Through a combination of computational modeling and a pilot study of actual hardware in cadaveric temporal bone specimens, we demonstrated that the high resolution and additional distal degree-of-freedom that our scope provides enables physicians to capture high quality images and videos of the middle ear space using a trans-ET approach for diagnostic purposes. Our computational modeling work demonstrated that the use of a steerable endoscope grants more comprehensive visualization of the middle ear space when compared to a rigid scope, with high statistical significance. Further, a pilot study in cadaveric temporal bone specimens demonstrated that our steerable endoscope prototype achieves acceptable to excellent visualization of critical structures within the middle ear, as determined by an unbiased panel of otolaryngologists. These results, while preliminary, indicate significant promise in terms of reducing the invasiveness of second-look surgeries in the diagnosis and treatment of cholesteatoma and other ME disease. Current work is focused on developing a clinically-ready prototype and performing a series of test in whole-head cadaveric specimens, where we aim to quantitatively and statistically demonstrate trans-ET surveillance as a viable diagnostic option for middle ear pathologies.

Conflict of Interest

The authors have no conflicts of interest to report.

Acknowledgments

The authors would like to acknowledge Enable, Inc. for assisting in the design, development, and fabrication of the steerable endoscope presented in this work.

Financial Support: The authors thank the National Institutes of Health (NIH) for Grant R21 DC016153-01A1 which supported the work described in this paper. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Institutes of Health.

Role of Sponsors: The sponsors had no role in the design of the study, the collection and analysis of the data, or the preparation of the manuscript.

Statement of Ethical Approval: All temporal bone experiments were performed under the supervision of a single otolaryngologist (M.F.). All cadaver specimens were obtained from Vanderbilt University Medical Center, and their use falls under IRB Exemption according to 45 CFR 46. We did not recruit a multiple-user cohort to perform these pilot experiments. Recruitment of subjects for the purposes of Likert survey data collection was performed under the approved IRB protocol 200449. The computer simulations run on patient models were deemed exempt from the Worcester Polytechnic Institute IRB pursuant to 45 CFR 46.101(b). Based on these consideration, and the fact that no live patients or live animals were involved in our experiments, prior approval from an ethics committee (IRB or IACUC) was not required.

References

1. Ayache, D., V. Darrouzet, F. Dubrulle, C. Vincent, S. Bobin, M. Williams, and C. Martin. Imaging of non-operated cholesteatoma: Clinical practice guidelines. European Annals of Otorhinolaryngology, Head and Neck Diseases 129:148 – 152, 2012.

2. Barath, K., A. Huber, P. Stampfli, Z. Varga, and S. Kollias. Neuroradiology of Cholesteatomas. Am J Neuroradiol 32:221–229, 2011.

3. Bruns, T. L., K. E. Riojas, D. S. Ropella, M. S. Cavilla, A. J. Petruska, M. H. Freeman, R. F. Labadie, J. J. Abbott, and R. J. Webster. Magnetically steered robotic insertion of cochlear-implant electrode arrays: System integration and first-in-cadaver results. IEEE Robotics and Automation Letters 5:2240–2247, 2020.

4. Caversaccio, M., W. Wimmer, J. Anso, G. Mantokoudis, N. Gerber, C. Rathgeb, D. Schneider, J. Hermann, F. Wagner, O. Scheidegger, M. Huth, L. Anschuetz, M. Kompis, T. Williamson, B. Bell, K. Gavaghan, and S. Weber. Robotic middle ear access for cochlear implantation: First in man. PLOS ONE August:1–12, 2019.

5. Chang, P. and S. Kim. Cholesteatoma - Diagnosing the unsafe ear. Australian Family Physician 37:631–640, 2008.

6. Crowson, M. G., V. H. Ramprasad, N. Chapurin, C. D. Cunningham, and D. M. Kaylie. Cost analysis and outcomes of a second-look tympanoplasty-mastoidectomy strategy for cholesteatoma. Laryngoscope 126:2574–2579, 2016.

7. Dahroug, B., B. Tamadazte, S. Weber, L. Tavernier, and N. Andreff. Review on Otological Robotic Systems: Toward Microrobot-Assisted Cholesteatoma Surgery. IEEE Reviews in Biomedical Engineering 11:125–142, 2018.

8. Di Martino, E., L. E. Walther, and M. Westhofen. Endoscopic Examination of the Eustachian Tube: A Step-by-Step Approach. Otology and Neurology 26:1112–1117, 2005.

9. Dillon, N. P., R. Balachandran, J. M. Fitzpatrick, M. A. Siebold, R. F. Labadie, G. B. Wanna, T. J. Withrow, and I. Webster, Robert J. A Compact, Bone-Attached Robot for Mastoidectomy. Journal of Medical Devices 9, 2015. 031003.

10. DTIC. MIL-STD-1472D: Section on Rotary Controls , 1989.

11. Fedorov, A., R. Beichel, J. Kalpathy-Cramer, J. Finet, J.-C. Fillion-Robin, S. Pujol, C. Bauer, D. Jennings, F. Fennessy, M. Sonka, J. Buatti, S. Aylward, J. V. Miller, S. Pieper, and R. Kikinis. 3d slicer as an image computing platform for the quantitative imaging network. Magnetic Resonance Imaging 30:1323 – 1341, 2012. Quantitative Imaging in Cancer.

12. Fichera, L., N. P. Dillon, D. Zhang, I. S. Godage, M. A. Siebold, B. I. Hartley, J. H. Noble, P. T. Russell, R. F. Labadie, and R. J. Webster. Through the eustachian tube and beyond: A new miniature robotic endoscope to see into the middle ear. IEEE robotics and automation letters 2:1488–1494, 2017.

13. Gulati, M., S. Gupta, A. Prakash, A. Garg, and R. Dixit. HRCT imaging of acquired cholesteatoma : a pictorial review. Insights Into Imaging 10:4–11, 2019.

14. Haginomori, S.-I., A. Takamaki, R. Nonaka, and H. Takenaka. Residual Cholesteatoma: Incidence and Location in Canal Wall Down Tympanoplasty with Soft-Wall Reconstruction. Archives of Otolaryngology - Head and Neck Surgery 134:652–657, 2008.

15. Hannula, S., R. Bloigu, K. Majamaa, M. Sorri, and E. Mäki-Torkko. Ear diseases and other risk factors for hearing impairment among adults : An epidemiological study Ear diseases and other risk factors for hearing impairment among adults : An epidemiological study. International Journal of Audiology 51:833–840, 2012.

16. Heywood, R. L. and A. A. Narula. The pros and cons of canal wall up versus canal wall down mastoidectomy for cholesteatoma. Otorhinolaryngologist 6:140–143, 2013.

17. Ho, S. and J. Kveton. Efficacy of the 2-Staged Procedure in the Management of Cholesteatoma. Archives of Otolaryngology - Head and Neck Surgery 129:541–545, 2003.

18. Jaiprakash, A., W. B. O'Callaghan, S. L. Whitehouse, A. Pandey, L. Wu, J. Roberts, and R. W. Crawford. Orthopaedic surgeon attitudes towards current limitations and the potential for robotic and technological innovation in arthroscopic surgery. Journal of Orthopaedic Surgery 25:1–6, 2017.

19. Karhuketo, T. S. and H. J. Puhakka. Middle Ear Imaging via the Eustachian Tube with a Superfine Fiberoptic Videomicroendoscope. ORL 60:30–34, 1998.

20. Keeler, J. A. and D. M. Kaylie. Cholesteatoma: Is a Second Stage Necessary? The Laryngoscope 126:1499–1500, 2016.

21. Klug, C., B. Fabinyi, and M. Tschabitscher. Endoscopy of the Middle Ear Through the Eustachian Tube: Anatomic Possibilities and Limitations. The American Journal of Otology 20:209–303, 1999.

22. Kuo, C.-l., A.-s. Shiao, M. Yung, M. Sakagami, H. Sudhoff, C.-h. Wang, C.-h. Hsu, and C.-f. Lien. Updates and Knowledge Gaps in Cholesteatoma Research. BioMed Research International 2015:1–16, 2015.

23. LaValle, S. M. Planning algorithms. Cambridge University Press , 2006.

24. Lingam, R. K., R. Nash, A. Majithia, A. Kalan, and A. Singh. Non-echoplanar diffusion weighted imaging in the detection of post-operative middle ear cholesteatoma : navigating beyond the pitfalls to find the pearl. Insights into Imaging 7:669–678, 2016.

25. Linstrom, C. J., C. A. Silverman, A. Rosen, and L. Z. Meiteles. Eustachian Tube Endoscopy in Patients With Chronic Ear Disease. The Laryngoscope 110:1884–1889, 2000.

26. Louw, L. Acquired cholesteatoma pathogenesis : stepwise explanations. The Journal of Laryngology and Otology 124:587–593, 2010.

27. Miroir, M., Y. Nguyen, O. Sterkers, and A. B. Grayeli. Design, Kinematic Optimization, and Evaluation of a Teleoperated System for Middle Ear Microsurgery. The Scientific World Journal 2012:1–19, 2012.

28. Möller, T. and B. Trumbore. Fast, minimum storage ray-triangle intersection. J. Graph. Tools 2:21–28, 1997.

29. Noble, J. H., B. M. Dawant, F. M. Warren, and R. F. Labadie. Automatic identification and 3D rendering of temporal bone anatomy. Otology and Neurotology 30:436–442, 2009.

30. Paltura, C., T. S. Can, B. K. Yilmaz, M. E. Dinç, Ö. N. Develiolu, and M. Külekçi. Eustachian tube diameter: Is it associated with chronic otitis media development? American Journal of Otolaryngology - Head and Neck Medicine and Surgery 38:414–416, 2017.

31. Plouin-Gaudon, I., D. Bossard, S. Ayari-Khalfallah, and P. Froehlich. Fusion of MRIs and CT Scans for Surgical Treatment of Cholesteatoma of the Middle Ear in Children. Archives of Otolaryngology - Head and Neck Surgery 136:878–883, 2010.

32. Poe, D., E. Rebeiz, M. Pankratov, and S. Shapshay. Transtympanic Endoscopy of the Middle Ear. The Laryngoscope 102:993–996, 1992.

33. Pusalkar, A. G. Cholesteatoma and Its Management. Indian Journal of Otolaryngology and Head & Neck Surgery 67:201–204, 2015.

34. Swaney, P. J., P. A. York, H. B. Gilbert, J. Burgner-Kahrs, and R. J. Webster. Design, fabrication, and testing of a needle-sized wrist for surgical instruments. Journal of medical devices 11:014501, 2017.

35. Swearingen, B. Update on Pituitary Surgery. The Journal of Clinical Endocrinology Metabolism 97:1073–1081, 2012.

36. Todt, I., R. Seidl, A. Ernst, I. Todt, R. Seidl, and A. Ernst. A new minimally invasive method for the transtubal, microendoscopic application of fluids to the middle ear appli-

cation of fluids to the middle ear. Minimally Invasive Therapy and Allied Technologies 17:300–302, 2009.

37. York, P. A., P. J. Swaney, H. B. Gilbert, and R. J. Webster. A wrist for needle-sized surgical robots. In: 2015 IEEE International Conference on Robotics and Automation (ICRA), pp. 1776–1781. 2015.

38. Zhang, D., M. L. Bennett, R. F. Labadie, and J. H. Noble. Simulation of trans-nasal endoscopy of the middle ear for visualization of cholesteatoma. In: 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI), pp. 1415–1418, IEEE2015.