Interactive Visualization for Image Guided Medialization Laryngoplasty

Can Kirmizibayrak*,‡, Mike Wakid*, James K. Hahn, Ph.D* and Steven Bielamowicz, MD†
*Department of Computer Science, The George Washington University, Washington, DC, USA
†Division of Otolaryngology, The George Washington University, Washington, DC, USA
‡email:kirmizi@gwmail.gwu.edu

***Abstract*—Surgical procedures require the physician to use preoperative and intra-operative information for deciding the optimal course of action to be taken on the patient. Medical visualization techniques are used to 'guide' the surgeon during the procedure to improve accuracy. We are presenting a visualization approach for image guided medialization laryngoplasty. A brief background is provided to describe the goal of the surgery. The problems associated with visualizing the available information and the proposed solutions are discussed. Our solutions aim to improve the mental registration of the virtual and real worlds by adding visual cues to the visualization and by interactive volume visualization and exploration schemes. Our approach displays medical images, image/video feeds and computer simulations in a single view. Methods were implemented on the GPU to achieve real time frame rates. Optical and electromagnetic tracking methods were used for direct interaction on the patient.**

***Index Terms*—Medical Visualization, User Interaction, Tracking, Magic Lens, Image Fusion**

I. Introduction

Image guidance has recently been and is continuing to be an integral part of surgical procedures. Image-guided surgery techniques can employ advanced visualization methods to 'guide' the surgeon during the surgery, to improve the understanding of available information and help the surgeon with the decision making process. In this paper, we present a visualization approach applied to a medical scenario, namely medialization laryngoplasty. An estimated 7.5 million people suffer from voice disorders in the United States [1]. Medialization laryngoplasty (ML) is a corrective surgery for vocal fold paralysis or paresis; which are debilitating conditions leading to difficulty in voice production because of an asymmetrical vocal fold structure. In this surgery, a patient-specific implant is placed to support the paralyzed side of vocal fold to restore the functionality and reestablish symmetry. The location and shape of the implant is important in producing the desired outcome, and the reported revision rate is as high as 24% even for experienced surgeons [2].

Main reasons for this high failure rate are twofold: first, speech production process is very intricate and small changes in implant size and location can lead to significant changes in outcome. Therefore, accurate placement of the implant is very critical. Second, the target region (the vocal fold) is not visible to the surgeon during the procedure, leading to the necessity to make educated guesses based on medical imagery (i.e. CT scans of the patient and laryngoscope video) and surgeon's domain knowledge.

In this paper, we present a visualization approach to alleviate these problems. Our fundamental goal is to improve the spatial awareness of the user and the overall understanding of the visualized datasets by enabling interaction on the patient and incorporating color information that is usually not available on volumetric medical datasets. The interaction is done via a volumetric magic lens approach, in which the location and orientation of the magic lens filters can be interactively changed by the user using optical or electromagnetic tracking. The color information from external or endoscopic cameras is fused with the volume visualization to help the mental registration process. In other words, we aim to provide an intuitive way for the surgeon to find the correspondences between the pre-operative medical imaging modalities and the patient during the surgery.

II. Related Work

Medical visualization is a widely researched area, motivated by the abundance of volumetric medical imaging modalities, which give the physicians the ability to see 'what lies inside' the patient [16]. Various visualization methods such as context-preserving selective transparency [6] or illustrative rendering styles [7] have been proposed, and have been applied to a number of image-guided surgical scenarios [20]. Multimodal visualization techniques were used to present all available information in an effective manner ([8],[9]). To our knowledge, there are no proposed visualization approaches for medialization laryngoplasty.

Magic lenses were introduced as spatial filter interfaces to locally modify the visual appearance of objects, to enhance data of interest or to suppress distracting information to visualize complex datasets ([10],[11]). They have been applied to various visualization problems such as geospatial datasets ([12],[13]). 3D applications include volumetric lenses [14], or optically-inspired distortion lenses [15].

User interaction is an important part of medical visualization. Adjusting rendering parameters around user specified points of interest was used in various research efforts, for example by using distance-based enhancements [17], attenuation [18] or combinations of fuzzy semantic rules [19]. Tracked surgical tools have been used for image guidance procedures [20], most

[1]This work was supported by the National Institutes of Health grant R01 DC007125 and 2009 Cosmos Young Scholars Award.

commonly in neurosurgery [21].

Augmented reality has been used to integrate real and virtual environments in medical settings ([22],[25]). A common limitation of these techniques is that the real image is augmented with the virtual image, therefore capabilities of volume rendering such as arbitrary viewpoints are not utilized. For instance, Higgins et al. [24] fused target sites extracted from CT data to bronchoscope videos. As an alternative, mixed reality approaches augment virtual scenes with real images. Using such a method, Dey et al. [23] texture mapped polygonal isosurfaces with 2D images obtained from neuro-endoscopes, with possible loss of information because of the iso-surface extraction process.

III. Challenges in Medialization Laryngoplasty

The main goal of the surgery is to place a structural support into a small window cut on the laryngeal cartilage. In the current practice, the shape and location of the implant is planned before surgery after examinations of the patient, CT scans and a pre-operative laryngoscope (a flexible endoscope to observe the airway) study. Computer simulations of phonation process with patient-specific data [26] are also available. During surgery, all these datasets are considered and the implant is placed at a location as close as possible to the planned scenario. Therefore, a good visualization system should display all available information to improve the accuracy of the implant placement process.

One important aspect of such systems that is usually overlooked is the fact that the purpose of the visualization is to help *guide* the actions performed *on* the patient. In other words, an approach is helpful when the correspondence between the visualization and the patient can be correctly (and preferably easily) established by the surgeon. Finding the correct registration between medical datasets and the patient is a challenging research problem, and various methods have been proposed to find preoperative dataset to patient registration for surgical applications [20], and specifically for medialization laryngoplasty [5]. The focus of this paper is, given the correct registration, how to visualize this relation to make the *mental registration* easier. We consider the problem as a two-way process: for the interaction point of view, the user is interacting directly with the patient. On the other hand, the visualization is updated interactively with these interactions, and is augmented with visual cues to make the connection to the patient easier. In the following sections, we will be describing each of these steps.

IV. Medical Volume Data and Image/Video Fusion

Medical image acquisition techniques capture images with many methods that differ from optical imaging (e.g. cameras). While this difference enables the very advantages they provide to capture information that the eye normally does not see, it makes it hard to establish the correspondence between medical images and real environments. Even though medical training helps the physicians recognize features specific to imaging modalities, augmenting medical visualization with visible color information would ease the mental registration. In our case, two sources of video information are available, external and internal. External video represents what the surgeon can see during the surgery (exposed laryngeal cartilage and the surrounding skin tissue). Internal video can show the vocal fold and airway. The relation between the patient, the pre-operative CT data and the video feeds can be difficult to establish. Our approach aims to augment medical volume datasets with color images/videos obtained from cameras to improve the mental registration. We use a volume rendering pipeline supplemented by mesh-based texture mapping. To find the correct texture color values, we have to find corresponding color values in the real world for each pixel rendered in the virtual scene. We do this by finding a matching virtual rendering for every corresponding real image. The camera is calibrated [4] to find the intrinsic parameters and the virtual image is modeled using these parameters and the extrinsic parameters provided by the tracker to have a virtual rendering that is very close to identical to the real scene (Figure 1). For every frame, each vertex in the current viewport is assigned the corresponding color value from the real image. As the camera position changes, different parts of the mesh are continuously 'painted' with camera color values. This is calculated completely on the GPU, achieving speed-ups of an order of magnitude compared to CPU implementation.

This approach alleviates some fundamental problems encountered in medialization laryngoplasty. During the surgery, occlusion is a major problem. From the outside, the vocal fold (the target structure) is not visible to the surgeon. Internally, the laryngoscope can show the vocal fold, but its relation

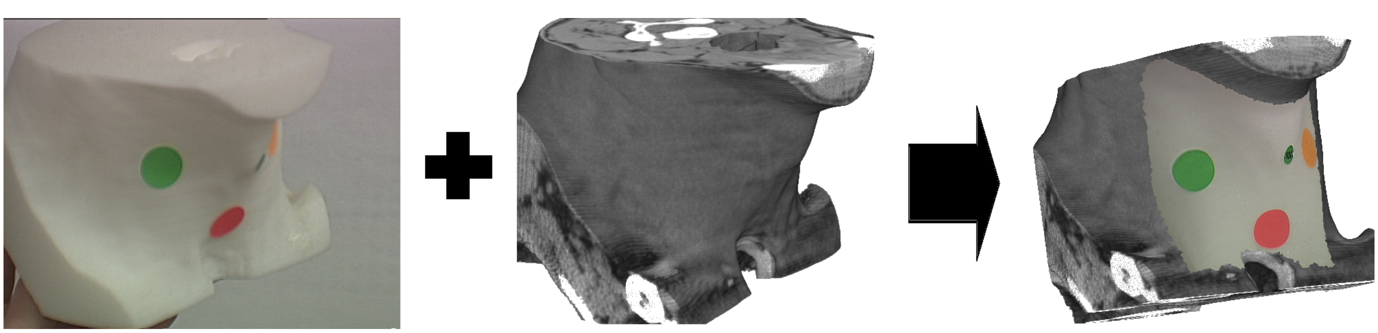

Fig. 1. Corresponding real (i.e. from camera, left) and virtual (i.e. volume rendered, center) views and image fusion result (right)

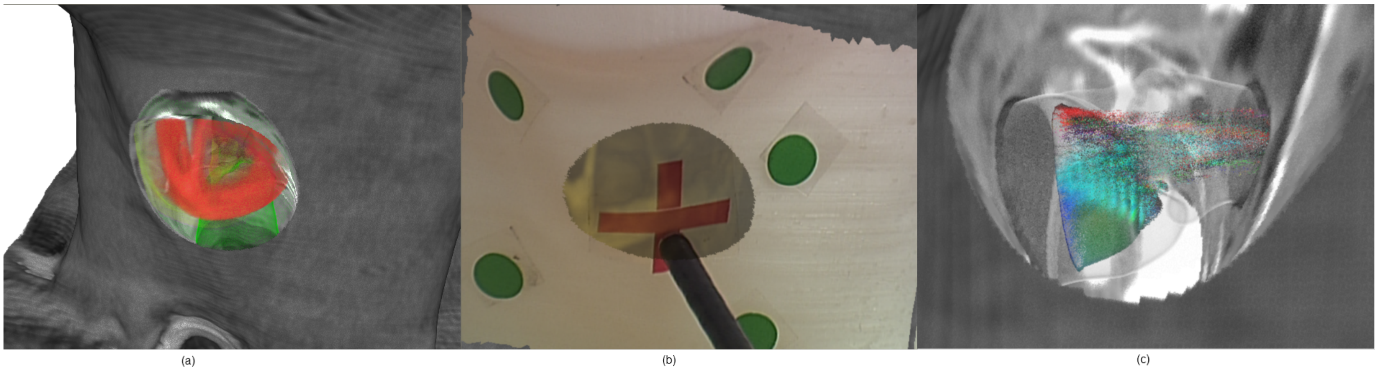

Fig. 2. (a) Magic Lens visualization with the combined use of two cylindrical lenses, (b) Selective transparency applied to a local region, (c) Flow visualization inside the transparent airway.

to structures visible from the outside is difficult to establish. Furthermore, navigation of the laryngoscope is difficult, limiting possible viewpoints. By fusing the camera images to the volume rendering, an intuitive connection to the real world can be realized, which can be viewed from any arbitrary viewpoint.

V. Interactive Magic Lens Visualization

During the medialization laryngoplasty procedure, the surgeon has multiple sources of information to consider, namely pre-operative images (e.g. CT scans), internal/external video and computer simulations. It is a challenging task to display all the data sources together because of inter- and intra-modality occlusion and possible cluttering of information. Our solution to this problem is a volumetric Magic Lens-inspired visualization approach to display all available data sources. The ‘lens’ in this method is a sub-volume the location of which is controlled by the user. The rendering style and data source inside each lens region (which will be referred to as a lens filter) can be selected by the user. Using real-time tracking, the user has the ability to select an ‘interesting’ part of the dataset by pointing to the corresponding region on the patient with a tracked stylus.

A. Pre-operative images

The most commonly used technique to display desired parts of a volume rendered dataset is the application of transfer functions. However, these functions are usually global therefore it is difficult to apply different transfer functions to selected local regions. By using a lens filter using different transfer functions, the user can change the rendering style in a local region by pointing the stylus to the desired region. By using combinations of lenses, complex transfer functions can be created by combining multiple lens filters. The shape and size of the lens can be arbitrary. Each of the filters can be assigned different transparency values to enable blending. In medialization larnygoplasty, the most important parts of the CT dataset are the laryngeal cartilage, the airway and the vocal fold. In Figure 2(a) two cylindrical lens filters are combined to show the cartilage (red) and the airway (green). The skin outside lens regions (gray) remains opaque and provides context to the lens filter regions.

B. Internal/external video

Fusing the color information to the visualization may occlude anatomical structures from the volume rendering. Adding semi-transparency alleviates this problem, but can result in loss of color values because of the ambiguity introduced by the blending. For this, we use a semitransparency lens filter to apply a user selected transparency value to a local region. The result of such a filter can be seen in Figure 2(b): the cartilage from the volume rendering is visible along with the mock incision target, while the rest of the fused color data is displayed opaquely to provide context.

C. Computer simulations

Computational Fluid Dynamics (CFD) simulations [26] are available to help design patient-specific implants for optimal surgical outcomes. Displaying this simulation results along with the anatomical information would help the surgeons understand the relation between the simulation data and patient anatomy. Our visualization system has a particle-based flow visualization lens filter to accomplish this task. Particles are advected using the velocity values from the CFD dataset. Particle advection process is done completely on the GPU using an approach similar to that briefly described in [3]. Each particle is assigned color/opacity values based on the corresponding CFD data value, such as pressure, velocity or vorticity. Figure 2(c) shows the transparent vocal fold and airway along with flow particles colored according to their vorticity values.

VI. Interaction

A crucial component of our visualization approach is enabling direct interaction with the patient. We believe that this helps the mental registration process by eliminating the abstraction in traditional interaction methods (e.g. mouse). We use tracking for interaction and have implemented two of the most widely used tracking techniques: electromagnetic (EM) and optical tracking. Tracking was also used for updating the camera position for the image/CT fusion process. Examples of tracking based interaction can be seen in Figure 3. After implementing both tracking methods for our application, we

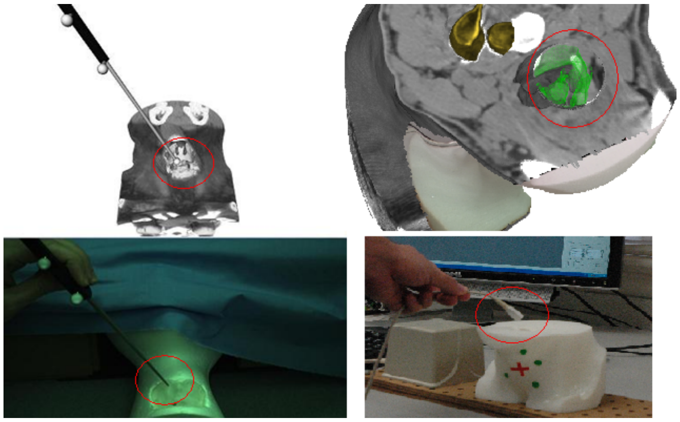

Fig. 3. Examples of two tracking techniques used: optical tracking (left), electromagnetic tracking (right). Bottom row shows real-time camera views while top row displays resulting visualization, with tracked tools and corresponding local visualization results circled in red.

have found EM tracking to be easier to use in a laboratory setting. The main reasons behind our preference are the lack of calibration and usability in minimally invasive settings.

VII. Conclusion and Future Work

In this paper, we have presented a visualization approach for image-guided medialization laryngoplasty that aims to improve the mental registration process to help the optimal implant placement. We aimed to reach this goal by providing visual cues from real world camera images and using an interactive Magic Lens approach to visualize multiple sources of information. By using features of modern GPUs, real-time refresh rates were achieved to provide immediate feedback to the surgeon. Our rendering algorithms were implemented on a Dell Precision 690 system with a NVIDIA Quadro FX 4600 graphics card with 768MB graphics memory. For our mock-up setup, we used a rapid prototyped (Dimension SST, Stratasys) phantom model built from a real-life patient CT scan to represent the patient in the real-life scenario. The rigid registration between the EM transmitter/optical tracker cameras and the phantom model was done manually. For tracking, we used 3D Guidance TrakStar from Ascension Corporation, Vicon MX-F40 cameras and a NDI Polaris system. The informal feedback we have received from surgeons that have used our system has been very positive. We are planning to undertake formal user studies to prove that our approaches help the mental registration process. Accuracy and validation studies to show if the implant placement accuracy improves would also be helpful to confirm our assumptions. We are planning to apply our approach to different medical scenarios, both for surgeries and for surgical planning.

References

[1] Statistics on Voice, Speech, and Language. http://www.nidcd.nih.gov/health/statistics/vsl.asp

[2] Anderson T.D., Spiegel J.R., and Sataloff R.T. Thyroplasty revisions: freq. and predictive factors. J Voice;17, :442-8, (2003).

[3] Li W., Fan Z., Wei X., Kaufman A.: Simulation with Complex Boundaries. In GPU Gems II, ed. Matt Pharr, Chapter 47. (2003).

[4] Tsai R. A versatile camera calibration technique for high-accuracy 3-D machine vision metrology using off-the-shelf TV cameras and lenses. IEEE J. of Robotics and Automation; 3(4): 323344, (1987).

[5] Jin G., Baek N., Hahn J.K., Bielamowicz S., Mittal R., Walsh R.: Image guided medialization laryngoplasty. In Computer Animation and Virtual Worlds 20, 1, 67-77. (2009).

[6] Kruger J., Schneider J., Westermann R.: ClearView: An Interactive Context Preserving Hotspot Visualization Technique. IEEE Trans. On Visualization and Computer Graphics 12, 5, 941-948. (2006).

[7] Svakhine, N., Ebert, D.S., Stredney, D.: Illustration motifs for effective medical volume illustration. IEEE Computer Graphics And Applications 25, 3, 31-39. (2005)

[8] Beyer, J., M. Hadwiger, et al. High-Quality Multimodal Volume Rendering for Preoperative Planning of Neurosurgical Interventions. Visualization and Computer Graphics, IEEE Transactions on 13(6): 1696-1703. (2007).

[9] Christian, R., R. Felix, et al. Interactive Visualization of Multimodal Volume Data for Neurosurgical Tumor Treatment. Computer Graphics Forum 27(3): 1055-1062. (2008).

[10] Bier, E. A., M. C. Stone, et al. Toolglass and magic lenses: the seethrough interface. Proc. of the 20th Ann. Conf. on Computer Graphics and Interactive Techniques. ACM: 73-80. (1993).

[11] Perlin, K. and D. Fox. Pad: an alternative approach to the computer interface. Proceedings of the 20th annual conference on Computer graphics and interactive techniques. Anaheim, CA, ACM: 57-64. (1993).

[12] Lee, S.-J., J. K. Hahn, et al. INSPECT: a dynamic visual query system for geospatial information exploration. Visualization and Data Analysis 2003, Santa Clara, CA, USA, SPIE. (2003).

[13] Borst, C. T., J; Best, C Real-Time Rendering Method and Performance Evaluation of Composable 3D Lenses for Interactive VR. Computer Graphics, IEEE Transactions on, vol. PP, no. 99, p. 1, (2009).

[14] Viega, J., M. J. Conway, et al. 3D magic lenses. Proc. of the 9th Ann. ACM Symp. on User interface software and technology. Seattle, Washington, United States, ACM. (1996).

[15] Wang, L., Y. Zhao, et al. The magic volume lens: an interactive focus+context technique for volume rendering. Visualization, 2005. VIS 05. IEEE. (2005).

[16] Correa, C. D. Visualizing what lies inside. SIGGRAPH Comput. Graph. 43(2): 1-6. (2009).

[17] Christian, R., R. Felix, et al. Interactive Visualization of Multimodal Volume Data for Neurosurgical Tumor Treatment. Computer Graphics Forum 27(3): 1055-1062. (2008).

[18] Marcelo, C., W. B. Ken, et al. The volume in focus: hardwareassisted focus and context effects for volume visualization. Proc. of the 2008 ACM Symp. on Applied computing., ACM. (2008).

[19] Rautek, P., S. Bruckner, et al. Interaction-Dependent Semantics for Illustrative Volume Rendering. Computer Graphics Forum 27(3): 847-854. (2008).

[20] Peters, T. M. Image-guidance for surgical procedures. Physics in Medicine and Biology(14): R505. (2006).

[21] Joshi, A., D. Scheinost, et al. Novel interaction techniques for neurosurgical planning and stereotactic navigation. Visualization and Computer Graphics, IEEE Trans. on 14(6): 1587-1594. (2008).

[22] Azuma R, Baillot Y, Behringer R, Feiner S, Julier S, MacIntyre B. Recent Advances in Augmented Reality. IEEE Computer Graphics and Applications ; 21(6): 34-47, DOI: 10.1109/38.963459. (2001)

[23] Dey, D, Gobbi DG, Slomka PJ, Surry KJ, Peters TM. Automatic fusion of freehand endoscopic brain images to three-dimensional surfaces: Creating stereoscopic panoramas. IEEE Transactions on Medical Imaging ; 21(1): 23-30, DOI: 10.1109/42.98123. (2002)

[24] Higgins WE, Helferty JP, Lu K, Merritt S, Rai L, Yu K. 3D CTvideo fusion for image-guided bronchoscopy. Computerized Medical Imaging and Graphics; 32(3): 159-173, (2008).

[25] Rolland, J. P. and H. Fuchs. Optical Versus Video See-Through Head-Mounted Displays in Medical Visualization. Presence: Teleoper. Virtual Environ. 9(3): 287-309. (2000).

[26] Luo H., Mittal R., Zheng X., Bielamowicz S., Walsh RJ., Hahn, JK.: An immersed-boundary method for flowstructure interaction in biological systems with application to phonation. Journal of Computational Physics 227, 22 9303-9332. (2008).