Accepted Manuscript online: 2024-07-02 Article published online: 2024-08-05

Development of a Machine Learning–Enabled Virtual Reality Tool for Preoperative Planning of Functional Endoscopic Sinus Surgery

Varun Gudapati[1] Alexander Chen[1] Scott Meyer[1] Chung-Chieh Jay Kuo[2] Yichen Ding[1] Tzung K. Hsiai[1] Marilene Wang[1]

[1] David Geffen School of Medicine, UCLA, Los Angeles, California, United States
[2] Ming-Hsieh Department of Electrical Engineering, USC, Los Angeles, California, United States

Address for correspondence Varun Gudapati, MD, David Geffen School of Medicine, UCLA, Los Angeles, CA 90095, United States (e-mail: VGudapati@mednet.ucla.edu).

J Neurol Surg Rep 2024;85:e118–e123.

Abstract

Objectives Virtual reality (VR) is an increasingly valuable teaching tool, but current simulators are not typically clinically scalable due to their reliance on inefficient manual segmentation. The objective of this project was to leverage a high-throughput and accurate machine learning method to automate data preparation for a patient-specific VR simulator used to explore preoperative sinus anatomy.

Methods An endoscopic VR simulator was designed in *Unity* to enable interactive exploration of sinus anatomy. The Saak transform, a data-efficient machine learning method, was adapted to accurately segment sinus computed tomography (CT) scans using minimal training data, and the resulting data were reconstructed into three-dimensional (3D) patient-specific models that could be explored in the simulator.

Results Using minimal training data, the Saak transform–based machine learning method offers accurate soft-tissue segmentation. When explored with an endoscope in the VR simulator, the anatomical models generated by the algorithm accurately capture key sinus structures and showcase patient-specific variability in anatomy.

Conclusion By offering an automatic means of preparing VR models from a patient's raw CT scans, this pipeline takes a key step toward clinical scalability. In addition to preoperative planning, this system also enables virtual endoscopy—a tool that is particularly useful in the COVID-19 era. As VR technology inevitably continues to develop, such a foundation will help ensure that future innovations remain clinically accessible.

Keywords
- computed tomography
- endoscopic sinus surgery
- imaging
- simulation training
- FESS

received March 5, 2024
accepted after revision May 16, 2024
accepted manuscript online July 2, 2024

DOI https://doi.org/10.1055/a-2358-8928.
ISSN 2193-6358.

Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

Introduction

The novel coronavirus disease 2019 (COVID-19) pandemic has thrust patients and health care workers into a vulnerable state. Faced with this global crisis, the medical field had to embrace new technologies to advance education and patient care.[1,2] Stay-at-home orders, widely implemented during the early phase of the pandemic, resulted in a dramatic decrease in the feasibility of in-person examinations. The subsequent increase in video and telehealth visits during this time period suggests the need for alternative, safer, and no-contact methods for examining patients, to avoid delays in diagnosis and treatment.

Virtual reality (VR) offers tremendous potential in the medical field, especially for inherently visual-spatial exercises like diagnostic and surgical endoscopy.[3–5] Sinus anatomy is intricate and variable, with close proximity to critical neurovascular structures.[6–8] Preoperative planning and innovative intraoperative image guidance systems presently rely on 2D computed tomography (CT) planes that may not offer the most intuitive visualization of anatomy.[9–12]

In otolaryngology, VR has demonstrated efficacy as a teaching tool for students, residents, and surgeons to hone procedural skills.[13–26] While such innovations showcase VR's potential, current simulators rely on laborious manual image segmentation—the identification of different components in an image—and are thus not clinically scalable.[20] Machine learning methods offer the potential to automate high-quality image segmentation, addressing a significant hurdle of clinical VR.[27–30]

Existing machine learning methods like convolutional neural networks (CNNs) have shown promise as a segmentation tool across a variety of modalities but require a large volume of high-quality annotated data, as observed in previous studies centered around sinus segmentation.[31–33] True clinical applicability demands a more data-efficient alternative.[34]

Subspace approximation with augmented kernels (Saak) is a novel transformation that offers a fully reversible and data-efficient means of feature extraction.[35,36] Equipping the Saak transform with a classifier produces an automatic image segmentation algorithm capable of operating with minimal training data. We previously developed this method and studied its ability to segment intricate light sheet fluorescence microscopy images, finding that Saak transform–based machine learning consistently outperformed a CNN, particularly with lower numbers of training images.[36]

In this study, we leverage data-efficient machine learning to create a VR tool for patient-specific surgical planning.[5] Our Saak transform–based method automatically segments soft tissue and bone from sinus CT scans to allow operators to explore a patient's unique anatomy in the VR domain.

Materials and Methods

Preparation of Training Data

All data collection for this study received IRB approval. We obtained Digital Imaging and Communications in Medicine (DICOM) files for three patients' sinus CT scans with identification stripped for confidentiality purposes. The patients were selected from a pool of preoperative candidates for functional endoscopic sinus surgery. Both the soft and bony tissues of all 548 axial images, between 100 and 200 per CT scan, were annotated in *Amira* to establish the ground truth for training and validation purposes.

Segmentation Algorithm

We used MATLAB to extract and trim each axial slice from the raw CT scan DICOM files to 350×350 windows encompassing the airspace structures. We then used randomly selected images and their manually segmented labels to train our Saak-based machine learning algorithm, consisting of a multistage Saak transform—based on principle component analysis (PCA)—and a random forest classifier. The model segmented all desired anatomic structures at once for each axial slice. We ran the trained model to segment each CT scan such that the training data did not overlap with testing data.

Validation of Segmentation

The Saak-based method was validated using randomly selected, nonoverlapping axial images from the CT datasets. We calculated the intersection over union (IOU) and dice similarity coefficient (DSC) of the segmentation results of our algorithm trained with three, six, and nine training images.

Virtual Reality Demo

After obtaining segmented data, we reconstructed the three-dimensional (3D) object in Amira 6.1 and generated, compressed, and exported a surface model to *Autodesk Maya* for scaling and smoothing. We developed our VR demo in *Unity* using models exported from *Maya*. We finalized all steps in *Maya* and *Unity* with educational licenses.

We used an Acer Windows Mixed Reality Headset (Acer) as the VR viewer and an educational-purpose license version of Unity 5.5 (Unity Technologies) as the development engine. In a new *Unity* project, we imported the patient's 3D reconstructed head and mounted a probe with a virtual endoscopic camera. We enabled user control of the probe using either Acer Windows Mixed Reality Controllers or keyboard inputs and computed its coordinate position for mapping to 2D slices of the CT scan. Finally, we designed a Unity canvas that contained windows for the probe's location and the endoscopic camera's display. The outline of our VR pipeline is outlined in ►**Fig. 1**.

Results

Automatic Segmentation

We compared the Saak transform segmentation results of soft tissue (►**Fig. 2**, *blue*) and bone (►**Fig. 2**, *red*) with the ground truth (►**Fig. 2**, *white*). Under the condition of three, six, and nine training images, the DSC of soft tissue was 0.94 ± 0.05, 0.96 ± 0.04, and 0.98 ± 0.01, respectively, while the IOU was 0.89 ± 0.09, 0.92 ± 0.07, and 0.97 ± 0.02, respectively. In comparison to soft-tissue segmentation, bone images had DSCs of 0.30 ± 0.06, 0.66 ± 0.10, and 0.6 ± 0.07

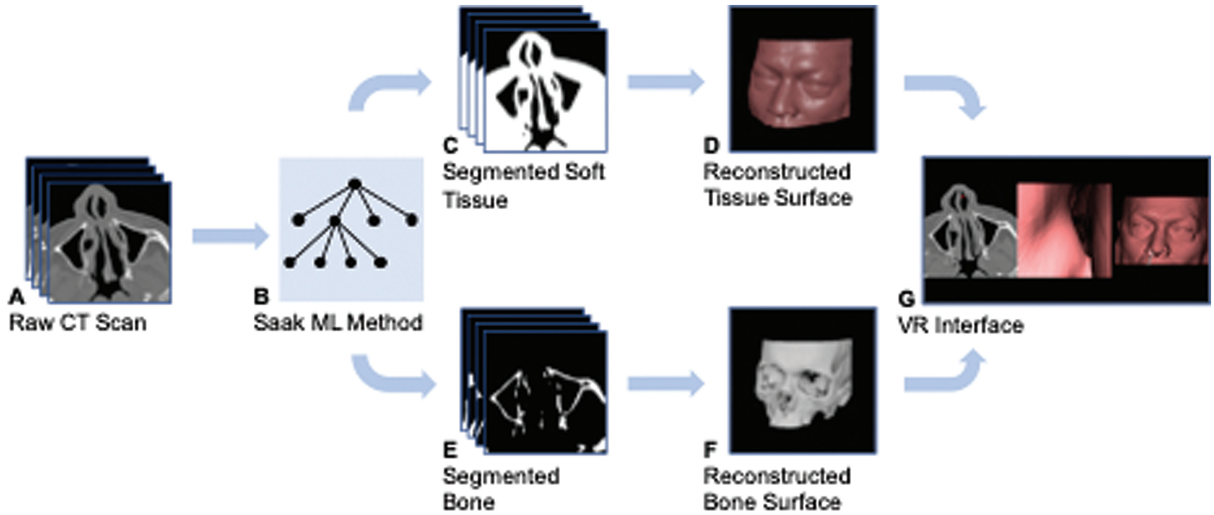

Fig. 1 Segmentation and virtual reality (VR) pipeline. (**A**) The axial slices of a raw computed tomography (CT) scan are passed to (**B**) the Saak transform–based machine learning algorithm, which has been trained with manually labeled images. The algorithm produces segmented slices of (**C**) soft tissue and (**E**) bone, which are stacked and processed to generate (**D, F**) three-dimensional (3D) meshes that can be ported to (**G**) the prebuilt *Unity* VR user interface for interactive anatomical exploration.

and IOUs of 0.44 ± 0.08, 0.49 ± 0.11, and 0.44 ± 0.08 across the same numbers of training images.

Virtual Reality Model

The automatic segmentation results enabled us to explore each patient's sinus anatomy in our *Unity* VR model with functionality to augment the user's experience. In the digital reconstruction, we examined the nares and nasal cavity to view the orifice of the maxillary sinus, ethmoid air cells, frontal sinus, nasopharynx, and any obstructions along the pathway with real-time mapping of our location (►**Fig. 3A, B**) to corresponding 2D slices of a CT scan. At any point during the exploration, we could toggle between soft-tissue and bone views to assess the degree of mucosal obstruction (►**Fig. 3C–F**). Our investigative trajectory was also traced during the virtual endoscopy, and tissue and bone boundaries were overlaid on the 3D path to assess the user's proximity to sensitive structures such as the lamina papyracea.

We performed virtual endoscopy on two patients with significant sinus disease and captured parallel views. In addition to the frontal sinus, we visualized the alar cartilage (►**Fig. 4A, E**), nasal cavity (►**Fig. 4B, F**), nasopharynx (►**Fig. 4C, G**), maxillary sinuses, and ethmoid air cells (►**Fig. 4D, H**). All structures were identifiable both through the navigation system and through the endoscopic camera's view. These perspectives allowed us to compare the obstructions and varying landmark locations between these patients (►**Fig. 4**).

Discussion

Both anatomic abnormalities and low conceptual expertise of the surgeon are cited as risk factors for increased complication rates in endoscopic sinus surgeries.[37] As VR enables a more intuitive, 3D visualization of anatomic features compared with traditional 2D CT scan views, our model has the potential to address both of these issues. Our framework offers the novel ability to automatically process and view a patient's unique anatomy in the VR domain. While existing VR models serve as teaching tools, the scalable patient-specific nature of our model broadens its application to preoperative planning, virtual endoscopy, and education. The mapping feature, inspired by the intraoperative image-guidance systems in practice today, further enhances the identification and understanding of landmarks.

The primary advantage of the Saak transform over other machine learning methods is its data efficiency. We were able to generate viable 3D sinus reconstructions using as few as three training images, meaning that users can tailor the algorithm to their specific segmentation needs with a minimal amount of manually labeled data. We assessed the accuracy of our Saak-based segmentation method with qualitative and quantitative measures. In addition to strong DSC and IOU values, the soft-tissue segmentation results had minimal visible noise, avoiding unnecessary surface vertices in the final model that would otherwise affect the performance of the VR demo. Bone segmentation was less accurate, likely due to the limitations of generating manually labeled training data around difficult-to-visualize structures like the ethmoid air cell walls. However, while highly precise soft-tissue segmentation is essential for an effective VR model, less accurate bone segmentation still enables clear visualization of contours needed to identify the probe's relative position. The six and nine training sizes were more effective in producing a bony anatomy model, but all training sizes provided a similarly effective soft-tissue VR experience.

This study provides a foundation for future innovations in the VR domain. Adding interactive functionality such as the ability to cut or debride tissue would build on this foundation and allow trial runs of a surgery.

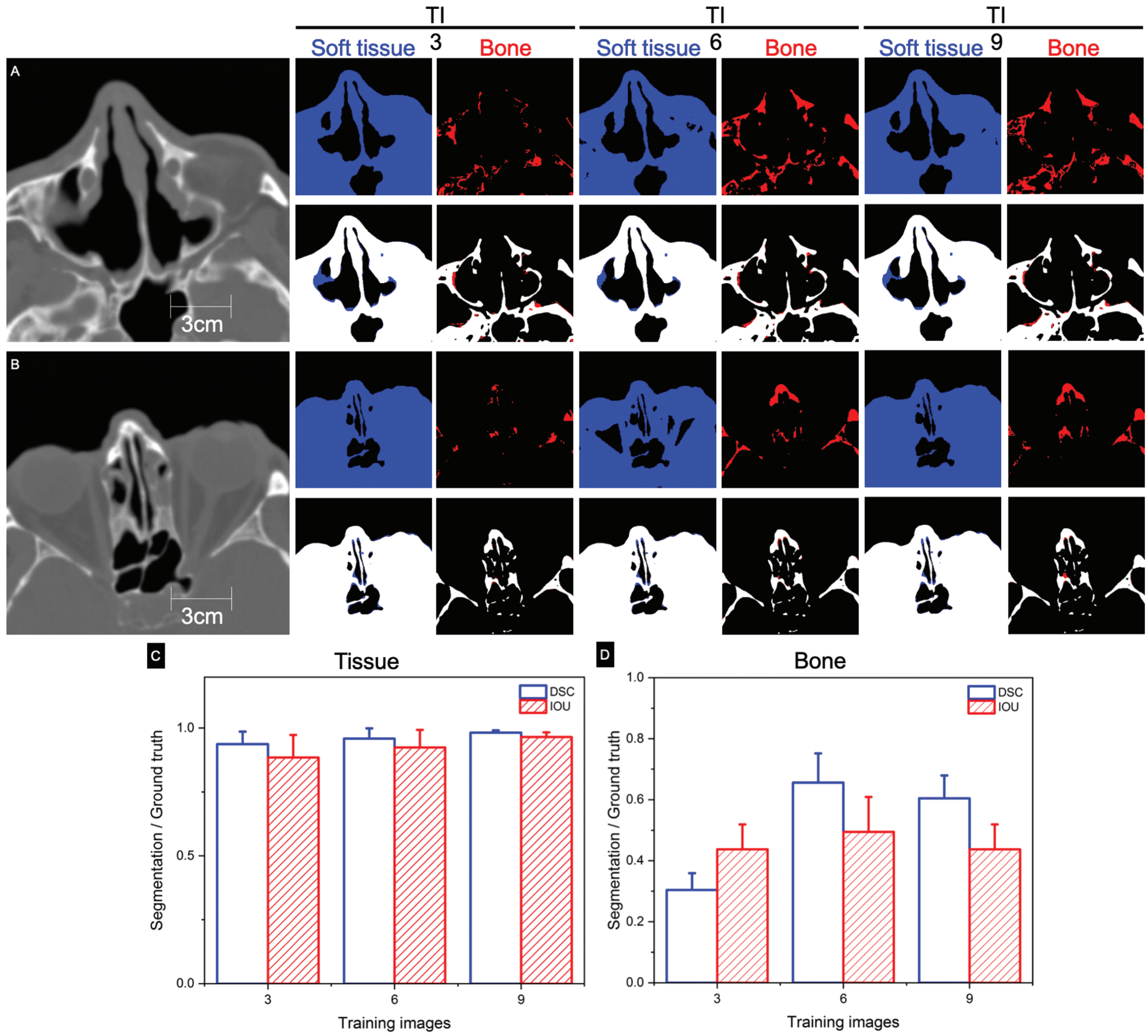

Fig. 2 Qualitative and quantitative comparison of segmentation results. (**A, B**) The Saak transform–based method's soft tissue (*blue*) and bone (*red*) segmentation results overlaid with the ground truth (*white*) for two axial slices of a computed tomography (CT) scan. (**C, D**) The dice similarity coefficient (DSC) and intersection over union (IOU) results of our automatic segmentation method for soft tissue and bone computed using 24 randomly selected validation image sets.

Both the machine learning and VR aspects of this study present limitations. Our Saak-based method functions well with consistent scan settings but is not designed to simultaneously handle images with multiple different windows and contrast profiles. However, the data-efficient nature of the Saak transform allows any user to tailor the performance of the algorithm for their scan standards using only minimal training data, preserving its clinical applicability.

Second, our VR model, like most others, is built using surface meshes rather than space-occupying voxels due to computational constraints. This makes deformation or manipulation of the object more challenging, limiting the realism of functional endoscopic sinus surgery simulators. Nevertheless, the fundamental framework of applying automatic segmentation to the VR domain is broadly applicable and will remain relevant even as voxel-based VR technology improves.

Conclusion

This study found that Saak transform–based machine learning automatically generates accurate, patient-specific VR models. Beyond preoperative planning, automatic segmentation and visualization of scans in VR may pave the way for virtual endoscopy and other remote alternatives to diagnostic examinations, addressing major challenges presented by the COVID-19 era. Future research into the automatic segmentation of additional anatomic structures and the interactive mechanics of VR will reinforce the clinical applicability of this technology.

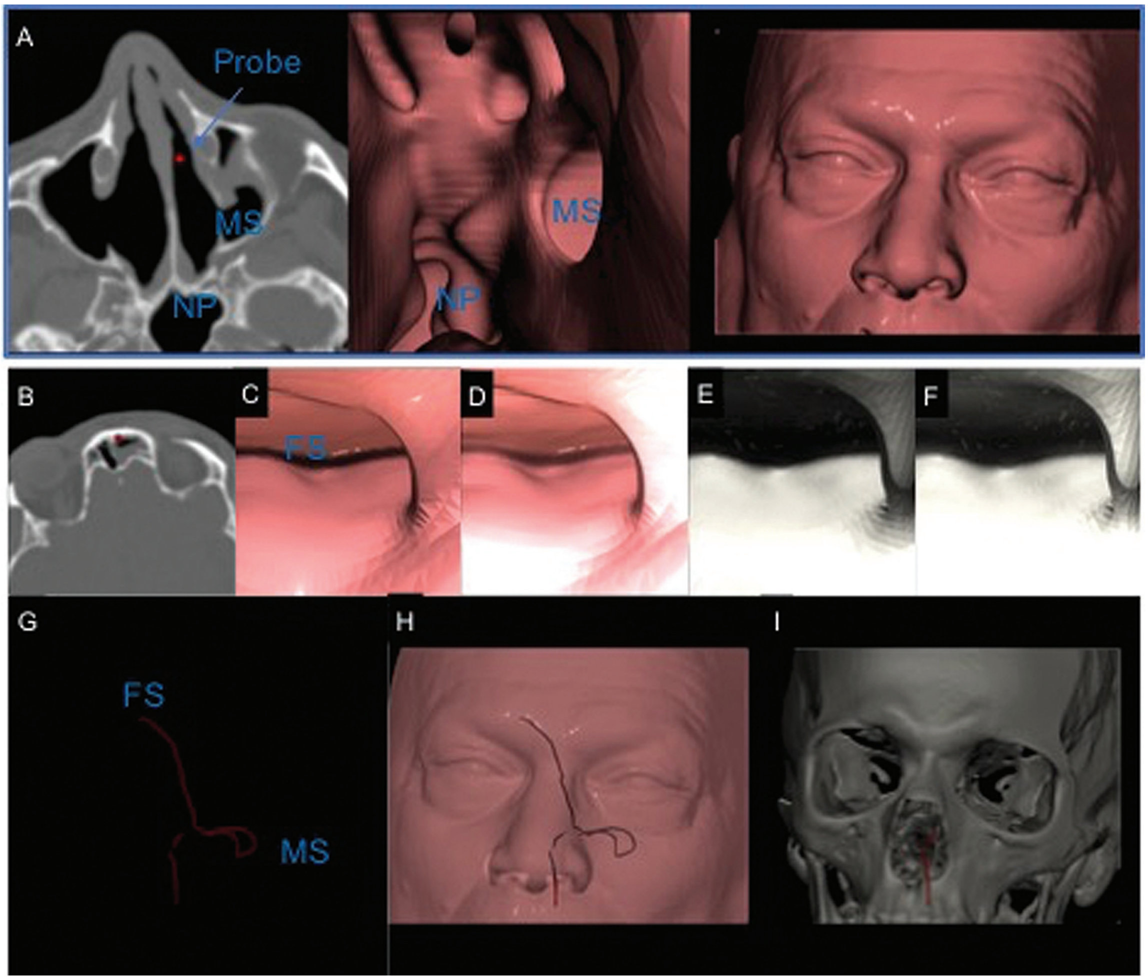

Fig. 3 Functionality of the virtual reality (VR) interface. (**A**) The three-paned Unity user interface displaying the mapped location of the probe, the probe's camera feed, and the three-dimensional (3D) model in which the probe is deployed (MS, maxillary sinus; NP, nasopharynx). (**B–F**) A view of the frontal sinus showing user ability to toggle between (**C, D**) tissue and (**E, F**) bone views and control the brightness of the probe's light. (**G–I**) A 3D tracer enables the user to view the probe's path. Bone and tissue can be toggled on and off to observe proximity of the path to other structures like the orbits.

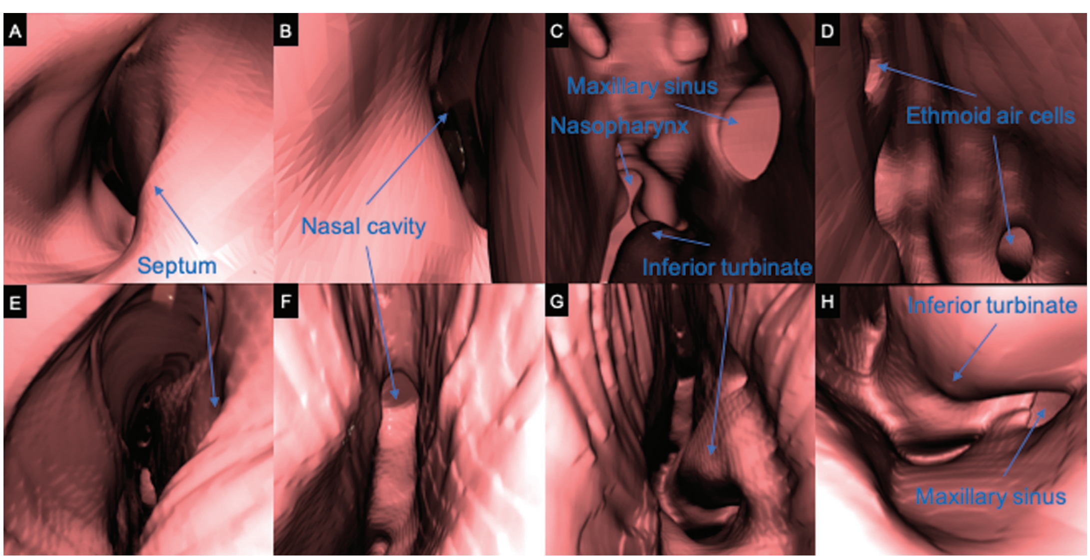

Fig. 4 Labeled anatomical features. (**A–D**) Selected views of a patient's anatomy. (**E–H**) Parallel views in the second patient highlighting the anatomical differences. (**B**) The first patient has a narrower nasal cavity due to obstruction compared with (**F**) the second patient. (**C**) While the natural orifice of the maxillary sinus is normally located above the inferior turbinate, (**H**) the second patient has a passageway below the turbinate from previous surgery.

Conflict of Interest

None declared.

References

1 Khariwala SS, Weinreich HM, McCoul ED, et al. Leveraging COVID-19-inspired changes to advance otolaryngology-here to stay. JAMA Otolaryngol Head Neck Surg 2020;146(07):605–607
2 Singh RP, Javaid M, Kataria R, Tyagi M, Haleem A, Suman R. Significant applications of virtual reality for COVID-19 pandemic. Diabetes Metab Syndr 2020;14(04):661–664
3 El Beheiry M, Doutreligne S, Caporal C, Ostertag C, Dahan M, Masson JB. Virtual reality: beyond visualization. J Mol Biol 2019; 431(07):1315–1321
4 Abiri A, Ding Y, Abiri P, et al. Simulating developmental cardiac morphology in virtual reality using a deformable image registration approach. Ann Biomed Eng 2018;46(12):2177–2188
5 Ding Y, Abiri A, Abiri P, et al. Integrating light-sheet imaging with virtual reality to recapitulate developmental cardiac mechanics. JCI Insight 2017;2(22):e97180
6 Mafee MF, Chow JM, Meyers R. Functional endoscopic sinus surgery: anatomy, CT screening, indications, and complications. AJR Am J Roentgenol 1993;160(04):735–744
7 Hosemann W. Comprehensive review on danger points, complications and medico-legal aspects in endoscopic sinus surgery. Laryngorhinootologie 2013;92(Suppl 1):S88–S136
8 Hosemann W, Grimm A. Surgical anatomy of the maxillary sinus. HNO 2020;68(08):555–565
9 Beswick DM, Ramakrishnan VR. The utility of image guidance in endoscopic sinus surgery: a narrative review. JAMA Otolaryngol Head Neck Surg 2020;146(03):286–290
10 Casiano RR, Numa WA Jr. Efficacy of computed tomographic image–guided endoscopic sinus surgery in residency training programs. Laryngoscope 2000;110(08):1277–1282
11 Fried MP, Parikh SR, Sadoughi B. Image-guidance for endoscopic sinus surgery. Laryngoscope 2008;118(07):1287–1292
12 Casiano RR. Intraoperative image-guidance technology. Arch Otolaryngol Head Neck Surg 1999;125(11):1275–1278
13 Rudman DT, Stredney D, Sessanna D, et al. Functional endoscopic sinus surgery training simulator. Laryngoscope 1998;108(11, Pt 1):1643–1647
14 Arora A, Lau LY, Awad Z, Darzi A, Singh A, Tolley N. Virtual reality simulation training in otolaryngology. Int J Surg 2014;12(02):87–94
15 Andersen SA, Konge L, Cayé-Thomasen P, Sørensen MS. Retention of mastoidectomy skills after virtual reality simulation training. JAMA Otolaryngol Head Neck Surg 2016;142(07):635–640
16 Barber SR, Jain S, Son YJ, Chang EH. Virtual functional endoscopic sinus surgery simulation with 3D-printed models for mixed-reality nasal endoscopy. Otolaryngol Head Neck Surg 2018;159(05):933–937
17 Kuppersmith RB, Johnston R, Jones SB, Jenkins HA. Virtual reality surgical simulation and otolaryngology. Arch Otolaryngol Head Neck Surg 1996;122(12):1297–1298
18 Fried MP, Satava R, Weghorst S, et al. Identifying and reducing errors with surgical simulation. Qual Saf Health Care 2004;13 (Suppl 1):i19–i26
19 Fried MP, Moharir VM, Shinmoto H, et al. Virtual laryngoscopy. Ann Otol Rhinol Laryngol 1999;108(03):221–226
20 Hsieh TY, Cervenka B, Dedhia R, Strong EB, Steele T. Assessment of a patient-specific, 3-dimensionally printed endoscopic sinus and skull base surgical model. JAMA Otolaryngol Head Neck Surg 2018;144(07):574–579
21 VanKoevering KK, Hollister SJ, Green GE. Advances in 3-dimensional printing in otolaryngology: a review. JAMA Otolaryngol Head Neck Surg 2017;143(02):178–183
22 VanKoevering KK, Malloy KM. Emerging role of three-dimensional printing in simulation in otolaryngology. Otolaryngol Clin North Am 2017;50(05):947–958
23 Hsieh TY, Cervenka B, Steele TO. Validation of haptic properties of materials for endoscopic sinus and skull base surgery simulation: reply. JAMA Otolaryngol Head Neck Surg 2018;144(12): 1185–1186
24 Gilani S, Norbash AM, Ringl H, Rubin GD, Napel S, Terris DJ. Virtual endoscopy of the paranasal sinuses using perspective volume rendered helical sinus computed tomography. Laryngoscope 1997;107(01):25–29
25 Gallivan RP, Nguyen TH, Armstrong WB. Head and neck computed tomography virtual endoscopy: evaluation of a new imaging technique. Laryngoscope 1999;109(10):1570–1579
26 Tolsdorff B, Pommert A, Höhne KH, et al. Virtual reality: a new paranasal sinus surgery simulator. Laryngoscope 2010;120(02): 420–426
27 Esteva A, Robicquet A, Ramsundar B, et al. A guide to deep learning in healthcare. Nat Med 2019;25(01):24–29
28 Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019;25(01):44–56
29 He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. Nat Med 2019;25(01):30–36
30 Panch T, Szolovits P, Atun R. Artificial intelligence, machine learning and health systems. J Glob Health 2018;8(02):020303
31 Humphries SM, Centeno JP, Notary AM, et al. Volumetric assessment of paranasal sinus opacification on computed tomography can be automated using a convolutional neural network. Int Forum Allergy Rhinol 2020;10(11):1218–1225
32 Iwamoto Y, Xiong K, Kitamura T, et al. Automatic segmentation of the paranasal sinus from computer tomography images using a probabilistic atlas and a fully convolutional network. Annu Int Conf IEEE Eng Med Biol Soc 2019;2019:2789–2792
33 Souadih K, Belaid A, Ben Salem D, Conze PH. Automatic forensic identification using 3D sphenoid sinus segmentation and deep characterization. Med Biol Eng Comput 2020;58(02): 291–306
34 Seo H, Badiei Khuzani M, Vasudevan V, et al. Machine learning techniques for biomedical image segmentation: an overview of technical aspects and introduction to state-of-art applications. Med Phys 2020;47(05):e148–e167
35 Kuo C-CJ, Chen Y. On data-driven Saak transform. J Vis Commun Image Represent 2018;50:237–246
36 Ding Y, Gudapati V, Lin R, et al. Saak transform-based machine learning for light-sheet imaging of cardiac trabeculation. IEEE Trans Biomed Eng 2021;68(01):225–235
37 Weber RK, Hosemann W. Comprehensive review on endonasal endoscopic sinus surgery. GMS Curr Top Otorhinolaryngol Head Neck Surg 2015;14:Doc08