High-Resolution Three-Dimensional Hybrid MRI + Low Dose CT Vocal Tract Modeling: A Cadaveric Pilot Study

*David MEYER, †Rushdi Zahid RUSHO, †Wahidul ALAM, ‡Gary E. CHRISTENSEN, §David M. HOWARD, #Jarron ATHA, #Eric A. HOFFMAN, ¶Brad STORY, ||Ingo R. TITZE, and #Sajan Goud LINGALA, *Winchester, Virginia, †‡#Iowa City, Iowa, §LDN, United Kingdom, ¶Tucson, Arizona, and ||Salt Lake City, Utah

Summary: Objectives. MRI based vocal tract models have many applications in voice research and education. These models do not adequately capture bony structures (e.g. teeth, mandible), and spatial resolution is often relatively low in order to minimize scanning time. Most MRI sequences achieve 3D vocal tract coverage at gross resolutions of 2 mm^3 within a scan time of <20 seconds. Computed tomography (CT) is well suited for vocal tract imaging, but is infrequently used due to the risk of ionizing radiation. In this cadaveric study, a single, extremely low-dose CT scan of the bony structures is blended with accelerated high-resolution (1 mm^3) MRI scans of the soft tissues, creating a high-resolution hybrid CT-MRI vocal tract model.

Methods. Minimum CT dosages were determined and a custom 16-channel airway receiver coil for accelerated high (1 mm^3) resolution MRI was evaluated. A rigid body landmark based partial volume registration scheme was then applied to the images, creating a hybrid CT-MRI model that was segmented in Slicer.

Results. Ultra-low dose CT produced images with sufficient quality to clearly visualize the bone, and exposed the cadaver to 0.06 mSv. This is comparable to atmospheric exposures during a round trip transatlantic flight. The custom 16-channel vocal tract coil produced acceptable image quality at 1 mm^3 resolution when reconstructed from ~6 fold undersampled data. High (1 mm^3) resolution MR imaging of short (<10 seconds) sustained sounds was achieved. The feasibility of hybrid CT-MRI vocal tract modeling was successfully demonstrated using the rigid body landmark based partial volume registration scheme. Segmentations of CT and hybrid CT-MRI images provided more detailed 3D representations of the vocal tract than 2 mm^3 MRI based segmentations.

Conclusions. The method described in this study indicates that high-resolution CT and MR image sets can be combined so that structures such as teeth and bone are accurately represented in vocal tract reconstructions. Such scans will aid learning and deepen understanding of anatomical features that relate to voice production, as well as furthering knowledge of the static and dynamic functioning of individual structures relating to voice production.

Keywords: Vocal tract−Airway−Modeling−Computed Tomography(CT)−Magnetic Resonance Imaging (MRI)−Segmentation.

INTRODUCTION

Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) have been used for 3D visualization of various anatomical structures in the head and neck[1,2] including the larynx.[3] In particular, vocal tract models derived from anatomical imagery have become important tools in voice research and education. They have been used to explore vocal tract area functions and voice quality,[4−8] physiological changes during vowel production,[9−11] and the effects of lowering fundamental frequency.[12] 3D printed models derived from MRI have also been shown to have similar acoustic properties as those found in vivo,[13] and freely downloadable model datasets have been shared online.[14]

In singing voice studies, MRI volumetric imaging and 3D printed models have been used to examine physiological changes when singers vary pitch and loudness,[15−17] vocal registers,[18−20] and to quantify the acoustic contribution of vallecullae[21] and piriform fossae.[22] They are a powerful educational tool for teaching the principles of vocal tract resonance and source/filter theory,[23] and have even been used to create a novel musical instrument, the Vocal Tract Organ.[24−26]

Problem

In comparison to many other techniques, MRI is an attractive modality for 3D vocal tract imaging because it poses little to no risk for participants. But because it is based on using the magnetic properties of hydrogen nuclei (which are abundant in tissues containing water and fat) to construct images of anatomical structures, it does not adequately capture the presence of the bony structures (e.g. teeth, mandible), thus compromising the accuracy of MRI modeled oropharyngeal spaces.[27] In addition, the duration of image acquisition is quite slow relative to the duration of most human sound production, and so the spatial resolution is often relatively low in order to minimize scanning time.

Thus, commercially available MRI sequences (e.g. those based on fast gradient echo) can typically achieve 3D vocal

Accepted for publication September 13, 2022.

From the *Janette Ogg Voice Research Center, Shenandoah University, Winchester, Virginia; †Roy J Carver Department of Biomedical Engineering, University of Iowa, Iowa City, Iowa; ‡Department of Electrical and Computer Engineering, University of Iowa, Iowa City, Iowa; §Department of Electronic Engineering, Royal Holloway, University of London, LDN, United Kingdom; #Departments of Radiology, Medicine and Roy J Carver Department of Biomedical Engineering, University of Iowa, Iowa City, Iowa; ¶College of Science, University of Arizona, Tucson, Arizona; and the ||Utah Center for Vocology, The University of Utah, Salt Lake City, Utah.

Address correspondence and reprint request to: David Meyer, Shenandoah University, Janette Ogg Voice Research Center, 1460 University Ave., Winchester, VA 22601 E-mail: dmeyer2@su.edu

Journal of Voice, Vol. ■■, No. ■■, pp. ■■−■■
0892-1997

https://doi.org/10.1016/j.jvoice.2022.09.013

tract coverage at gross resolutions of 2 mm^3 within a scan time of <20 seconds, which restricts voice tasks to a single exhalation.[28–30] At this resolution, modeling fine structures such as the epiglottis and vocal folds are compromised. While higher resolutions (e.g. < 1 mm^3) are desired, they require scan times of 75 seconds to 136 minutes,[31,32] which is, of course, far too long for sustained vocalization. Furthermore, commercially available receiver coils such as head only coil, head and neck coil, or neurovascular coils do not provide sufficient SNR over the entire vocal tract.[33] They show significant drop in sensitivity in the inferior regions of vocal tract (e.g. base of the tongue, glottis, and larynx). Accelerated MRI methods such as those based on parallel imaging with custom airway coils or constrained reconstruction can address the above tradeoffs,[34,35] and in this work we explore such methods to improve vocal tract reconstruction based on 3 Tesla MRI.

In contrast to MRI, computed tomography (CT) is well suited for acquisition of bony structures and air-tissue boundaries[36–39] in images, but is infrequently used for vocal tract studies due to the risk of ionizing radiation. Advances in medical imaging have led to substantially reduced CT dosages.[40–43] However, even with ultra-low dose CT, when sequential imaging is sought, the risk of ionizing radiation inherent in CT remains and must be mitigated. Determining minimum CT dosages for blended CT-MRI vocal tract model generation is a logical first step.

With the risk of ionizing radiation successfully mitigated, a single, extremely low-dose CT scan of the bony structures (which do not significantly change over time) could be blended with high-resolution MRI scans of the soft tissues, a hybrid CT-MRI vocal tract model may be constructed. Such a model would more accurately represent the structures used in speech and singing.

METHODS

This exploratory study used a recently deceased (<24 hours) unembalmed donor cadaver from the University of Iowa Deeded Body Program to explore the following:

1. The minimum CT dosage in milliSieverts (mSv) needed to maintain clear bone contours and air-tissue boundaries.
2. The evaluation of a custom 16-channel airway receiver coil for accelerated high (1 mm^3) resolution imaging of the vocal tract within a scan time of <10 seconds, and
3. The feasibility of blending CT and MR scans to construct a high-resolution hybrid model of the vocal tract.

1) CT dosage determination: The donor cadaver was scanned in the Advanced Pulmonary Physiologic Imaging Laboratory at the University of Iowa with a Siemens SOMATOM Force. Dose modulation was enabled, and various quality reference mAs's (QRM) were compared. QRM allows the mAs to fluctuate slice-by-slice based on the thickness of the structure being imaged. As QRM decreases, so does effective radiation exposure, measured in milliSieverts (mSv). All images were reconstructed using a Qr40 kernel, an ADMIRE (advanced model based iterative reconstruction[44]) strength of 5, Slice thickness and spacing of 1 mm (matching MR resolution), a pitch of 1.0, scan rotation time of 0.25 second, kVp of 120.

2) MR imaging was performed on a GE 3 Tesla Premier scanner equipped with high performance gradients (80 mT/m - maximum gradient strength, 200 mT/m/sec - slew rate). A gradient echo 3D sequence was used to scan the cadaver's upper-airway with a custom 16-channel upper-airway receiver coil using the following parameters: field of view: 24 cm × 24 cm × 12 cm; resolution: 1 mm × 1 mm × 1 mm; TR = 5.5 ms/TE = 2.5 ms; flip angle = 5 degrees, receiver bandwidth = 244.14 Hz/pixel; scan time = 77.4 secs. The scan was repeated with the same parameters but for a lower resolution: 2 mm × 2 mm × 2 mm, and shorter scan time of 19.44 seconds. The raw multi coil k-space data was saved and exported for offline processing. Coil sensitivity maps were estimated from the central 5% of k-space measurements using the sum of squares approach. Next, multiple coil volumes were reconstructed by applying two dimensional inverse Fourier Transforms (2D-IFT) on the acquired multi-coil k-space measurements. A single ground truth volume was formed by multiplying the complex conjugate of the coil maps to the multi coil volumes, followed by addition along the coil dimension. To evaluate acceleration capabilities of the airway coil, retrospective under-sampling was performed, in which the multi coil k-space data was sub-sampled by a variable density random order sampling scheme in the sagittal (ky-kz) plane. The acceleration factor (R) is defined as the total number of k-space samples in the ky-kz plane in the fully sampled data divided by the number of k-space samples in the sub-sampled dataset. The under sampled k-space data was then passed through an image domain based sparse sensitivity encoding (sparse-SENSE) parallel imaging constrained reconstruction algorithm to reconstruct the underlying volume from limited k-space samples. The algorithm exploits the diversity of receiver coil sensitivity profiles across the vocal tract, and sparse spatial finite differences of the images using a total variation constraint. We evaluated image quality using the Peak Signal to Noise ratio (PSNR) metric that captures the pixel-by-pixel error between the reconstructed images and the fully sampled reference image.

3) A rigid body landmark based partial volume registration scheme was used to match the MRI and CT volumes using corresponding manually identified anatomical landmarks in the MRI and CT modalities. These landmarks were: tip of the epiglottis, posterior boundary of inferior nasal concha, hyoid bone body midline, apex (tip) of the tongue, buccal portion of the bony ridge between left first and second premolars, and base of the mandible. The estimated transformation was used to transform the MR image into the coordinate system of the CT image. The bones and

the teeth were segmented from the CT image using intensity-based thresholding. These segmentations were then inserted into the registered MR image to create a hybrid image that had soft tissue imaged by MR, and the bones and teeth imaged by CT.

Vocal tract data from 2 mm^3 MR, 1 mm^3 CT, and the blended 1 mm^3 CT-MR images were imported into Slicer 5.1.0 r31082 / 115e32b as Digital Imaging and Communications in Medicine (DICOM) stacks. These were segmented to produce 3D models that were compared and visually evaluated by a subset of the authors (Meyer, Lingala, Howard) for anatomical accuracy. Manual slice-by-slice segmentation, "grow from seeds," and "threshold paint" methods were employed to ensure accurate reconstructions.

RESULTS

1) Image quality was evaluated to determine the minimum dosage needed to maintain clear bone contours and air-tissue boundaries. The lowest dose (QRM 5) produced images with sufficient quality to clearly visualize the bone, and represents a substantially mitigated risk for living participants in future studies (Figure 1) and allows for multiple scans or higher spatial resolution (thinner slices) should that be needed. Emerging photon counting CT technologies will provide for further reductions in ionized radiation, and higher spatial resolution.[43]

The following equation was used to calculate dose: mSv = DLP (mGy*cm) × k-factor, where the k-factor of the neck is 0.0059, and the k-factor of the head/neck is 0.0031. The two QRM-5 equations are as follows:

Head/neck QRM-5: 20.6 mGy*cm × 0.0031 = 0.06 mSv.
Neck QRM-5: 15.8 mGy*cm × 0.0059 = 0.09 mSv

Using this equation, the following dosages were calculated:

Neck QRM-10: 0.17 mSv
Neck QRM-5: 0.09 mSv
Head/Neck QRM-5: 0.06 mSv

By comparison, a single chest x-ray is on the order of 0.1 mSv,[45] and a New York to London round trip flight exposes travelers to 0.064 mSv.[46]

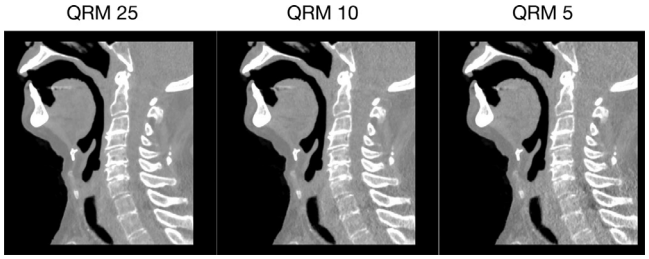

FIGURE 1. CT image evaluation at various QRMs.

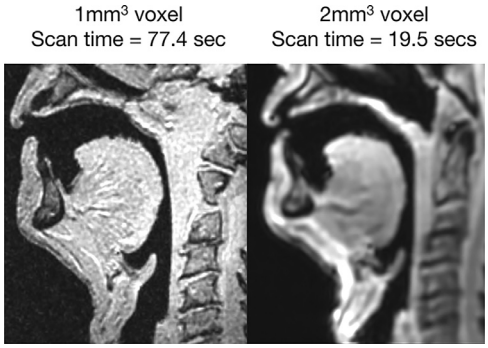

FIGURE 2. MR image at two resolutions.

2) At full sampling (i.e. Rnet = 1), MR scans at 1 mm^3 and 2 mm^3 voxel resolution were compared. Scans at 1 mm^3 voxel resolution displayed clearer air-tissue boundaries, reduced blurring, and were informally assessed by a subset of the authors (Meyer, Lingala, and Howard) to be superior in quality to images scanned at 2 mm^3 voxel resolution.

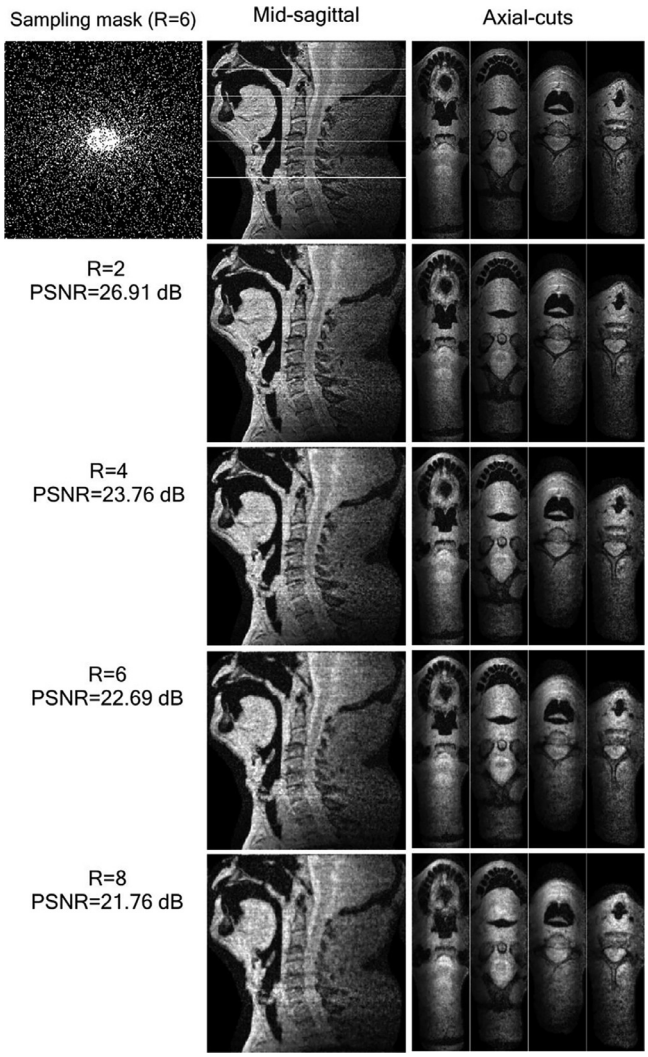

FIGURE 3. sparse SENSE reconstructions of 1 mm^3 data.

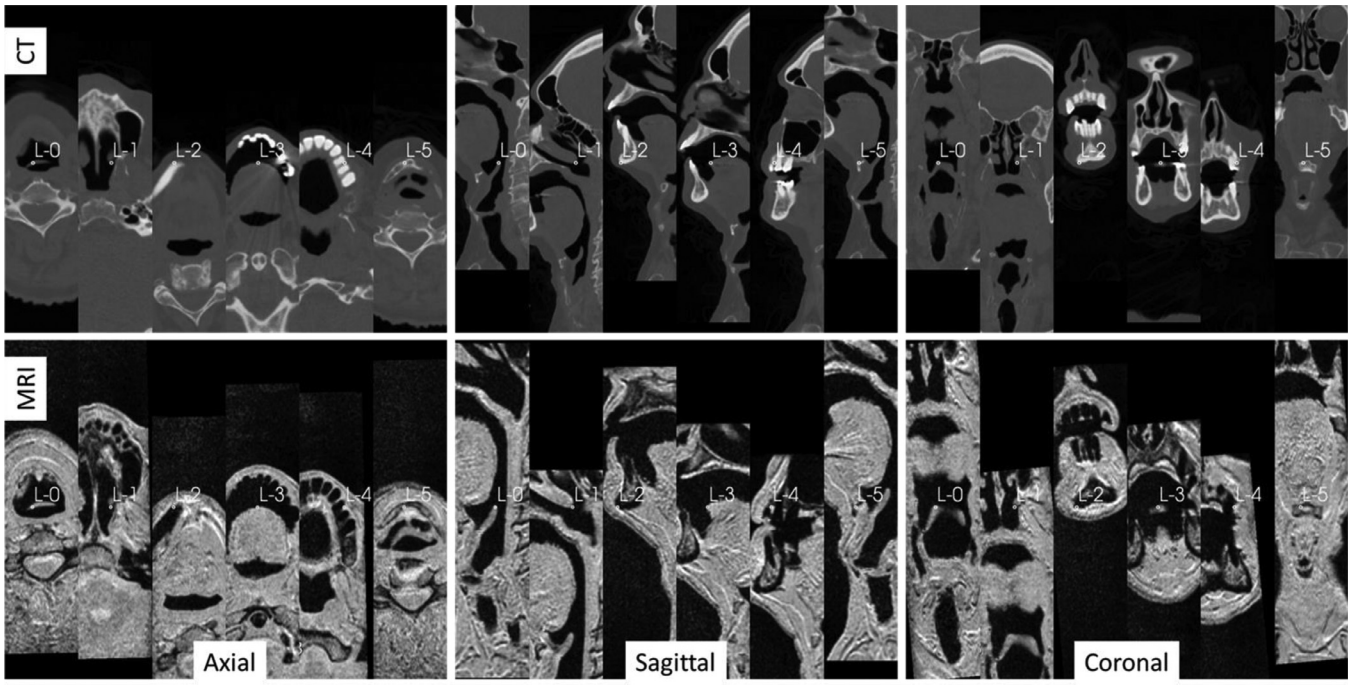

FIGURE 4. Landmark registration.

1 mm^3 and 2 mm^3 voxel resolution images are shown in Figure 2.

Figure 3 shows sparse SENSE reconstructions of the 1 mm^3 data at reduction. We observe robust reconstructions as a function of reduction factor (R). An example of the sampling mask used at R = 6 is also shown. The reconstructions gracefully degrade with increasing R. At R = 8 fold, there is still clear depiction of the various air-tissue boundaries, and image texture/features. The 8-fold gain in acceleration means that the scanning time of 1 mm^3 3D vocal tract MRI can be reduced by a factor of 8, which equals to ~9.6 seconds. The <10 seconds scan time can allow for high (1 mm^3) resolution imaging of short sustained sounds.

3) The rigid body landmark based partial volume registration scheme was successfully applied to match the manually identified anatomical landmarks in both the MRI and CT modalities. Figure 4 shows the locations of the landmarks in axial, sagittal, and coronal views for both MRI and CT images. The alignment of landmark registration from the MR and CT images may be seen in Figure 5, and when deemed acceptable, the MR image was transformed into the coordinate system of the CT image to create

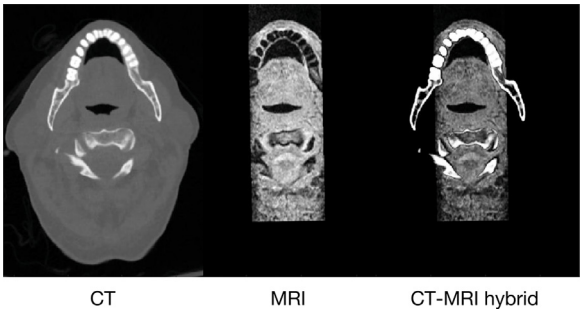

FIGURE 6. CT, MRI, and hybrid image-axial view.

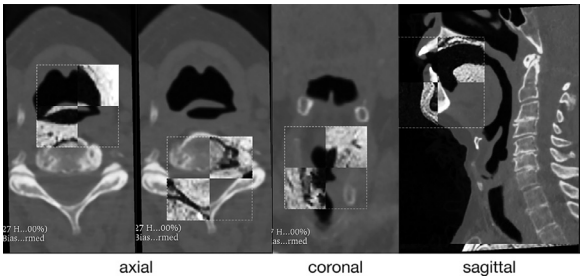

FIGURE 5. Registration evaluation.

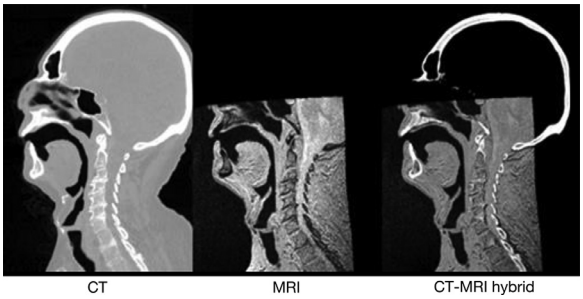

FIGURE 7. CT, MRI, and hybrid image-sagittal view.

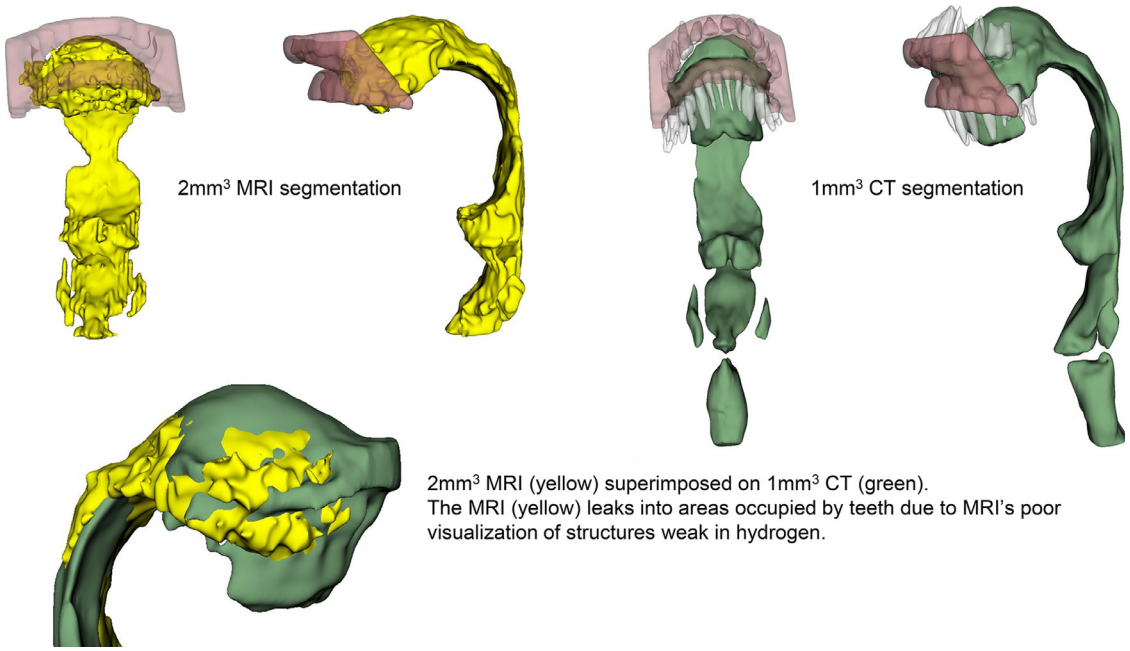

FIGURE 8. 2 mm^3 MRI (yellow) and 1 mm^3 CT (green) segmentation.

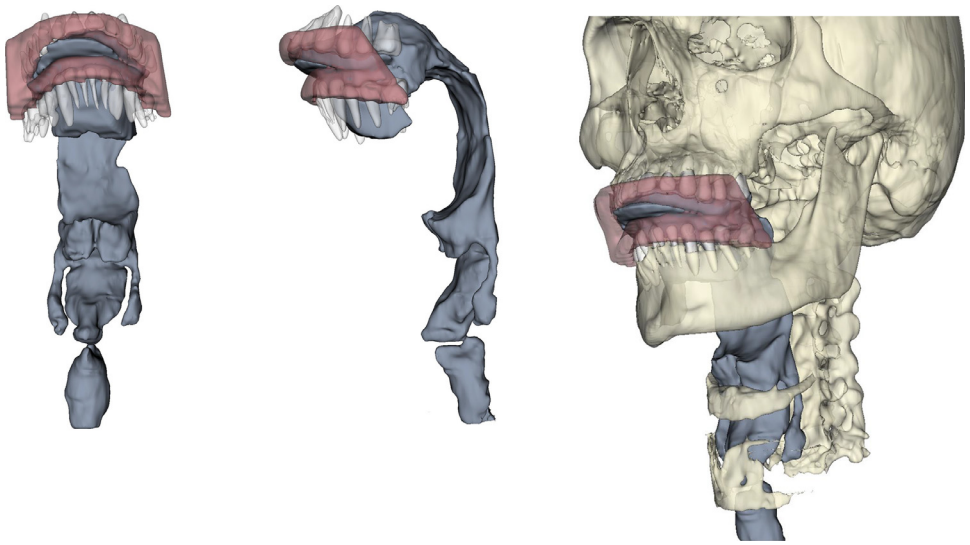

FIGURE 9. Hybrid 1 mm^3 CT-MRI (blue) segmentation.

a hybrid image that had soft tissue imaged by MR, and the bones and teeth imaged by CT (Figures 6-7).

2 mm^3 MRI, 1 mm^3 CT, and 1 mm^3 hybrid CT-MRI DICOM stacks were segmented in Slicer 5.1.0 r31082 to produce 3D models of the vocal tract (Figures 8 and 9). These 3D models were compared and visually evaluated by a subset of the authors (Meyer, Lingala, Howard) for anatomical accuracy. Figure 8 also includes a superimposed segmentation of the oral cavity, combining 2 mm^3 MRI (in yellow) with 1 mm^3 CT (in green). As stated earlier, as opposed to CT, MRI poorly images teeth. Therefore the superimposed image demonstrates the MRI segmentation (yellow) "leaking" into dental spaces that are more accurately imaged in CT (green).

DISCUSSION

The aims of this proof-of-concept investigation were: 1) to use an ultra-low dose CT protocol to establish minimum CT radiation dosage needed to maintain clear bone contours and air-tissue boundaries, 2) to evaluate a custom 16-channel airway receiver coil for accelerated high (1 mm^3) resolution imaging of the vocal tract, and 3) to determine the feasibility of low-dose hybrid CT-MRI vocal tract models with 1 mm^3 spatial resolution.

This method exposed the cadaver to very low levels of ionizing radiation, comparable to atmospheric exposures during a round trip transatlantic flight. This low level of exposure represents a substantially mitigated risk for living participants in future studies. Further reductions in ionized

radiation and higher spatial resolution may be possible with emerging photon counting CT technologies. The SOMATOM Force represents a current state-of-the-art CT technology with high spatial and contrast resolution along with current generation iterative reconstruction (ADMIRE) and a quantitative reconstruction kernel (Qr40).

As seen in Figure 3, imaging with the custom 16-channel vocal tract coil and finite difference reconstruction constraints produced acceptable image quality of the 1 mm^3 resolution data, when reconstructed from ~6 fold undersampled data. The image quality had an acceptable PSNR throughout the entire vocal tract. High (1 mm^3) resolution MR imaging of short (<10 seconds) sustained sounds was achieved.

The feasibility of hybrid CT-MRI vocal tract modeling was successfully demonstrated using the rigid body landmark based partial volume registration scheme. Segmentations produced from various imaging modalities displayed noteworthy characteristics (Figures 8 and 9). The CT and hybrid images appeared to provide more detailed representations of the vocal structures than the 2 mm^3 MR images. This was an expected finding due to the poor visualization of teeth in MRI,[27] and the complex geometry of the vocal tract's small structures. Details of the piriform sinuses and valleculae appeared to be best represented in the hybrid segmentation, and CT generally produced the clearest air-tissue boundaries.

LIMITATIONS AND FUTURE WORK

The techniques described in this study indicate that CT and MR image sets can be combined so that structures such as teeth and bone are accurately represented in vocal tract reconstructions. Still unknown, however, is whether the enhanced spatial accuracy of the airspace will translate into acoustic characteristics that are better aligned with those produced by an actual talker or singer than if MRI alone were used as the imaging modality.

Future work will include applying the CT/MRI hybrid technique to image sets collected from live talkers or singers (as opposed to cadavers) from which simultaneous audio recordings can also be obtained. Vocal tract reconstructions can then be processed to allow for various modeling and experimental approaches to be used to assess the acoustic characteristics, particularly the frequency response functions, relative to those measured from the talkers or singers.[47] These can include plane wave modeling, 2D and 3D finite-element modeling[48,49] as well as 3D printing to generate a physical structure than can be excited with a sound source to produce an output signal that can be analyzed.[14,50,22]

This method will also be applied in future work to more accurately image the mean vocal fold surface position and surface contour in the coronal view. Capturing such images is an ongoing challenge addressed in previous ultrasonic imaging studies[51,52] with mixed results.

CONCLUSION

The combined CT-MRI images enable segregation of a number of features based on how these are depicted best depending on their responses to CT or MRI scanning. This provides flexibility in image selection for users where different features can be explored by observing one or the other images (CT or MRI) or the combined rendering. Individual structures can be 3D printed to life-size to enable the vocal tract to be studied in detail and in the context of learning. It has been shown[53] that students using 3D anatomical models had advantages in test scores, accuracy, and student satisfaction over students in the 2D and cadaver groups. In addition, more students in the 3D printing group were satisfied with their learning compared with the conventional group.

Such scans will aid learning and deepen understanding of anatomical features that relate to voice production as well as furthering knowledge of the static and dynamic functioning of individual structures relating to voice production, extending the cylindrical vocal tract models of Arai,[54] relevant particularly for example to ENT and speech therapists, medical professionals, singers, orators, teachers, broadcasters, actors and their voice teachers. The possibility of handling and manipulating anatomical structures (in 3D printed form or within virtual and/or augmented reality) is a well-proven and rewarding way of gaining enhanced and more in-depth understanding of the function of component parts[55,56] as well as raising awareness in the public and future generations of researchers.[57]

Acknowledgements

The authors would like to thank Brian Chapman, Program Manager of the Deeded Body Program at the University of Iowa Carver College of Medicine for his assistance in this study. We are also profoundly grateful to those who in death have chosen to serve the living, dedicating their remains for use in science and education. We are honored by your gift. Thank you.

REFERENCES

1. Jones SB, Satava RM. Virtual endoscopy of the head and neck. Diagnosis using three-dimensional visualization and virtual representation. *Stud Health Technol Inform*. 1996;29:152–156
2. Hiramatsu H, Tokashiki R, Suzuki M. Usefulness of three-dimensional computed tomography of the larynx for evaluation of unilateral vocal fold paralysis before and after treatment: technique and clinical applications. *Eur Arch Otorhinolaryngol*. 2007;265:725–730. https://doi.org/10.1007/s00405-007-0514-7.
3. Rubin JS, Summers P, Harris T. Visualization of the human larynx: a three-dimensional computer modeling tool1. *Auris Nasus Larynx*. 1998;25:303–308. https://doi.org/10.1016/S0385-8146(98)00018-2.
4. Story BH, Titze IR, Hoffman EA. Vocal tract area functions from magnetic resonance imaging. *J Acoust Soc Am*. 1996;100:537–554. https://doi.org/10.1121/1.415960.
5. Story BH, Titze IR, Hoffman EA. The relationship of vocal tract shape to three voice qualities. *J Acoust Soc Am*. 2001;109:1651–1667. https://doi.org/10.1121/1.1352085.
6. Story BH. Comparison of magnetic resonance imaging-based vocal tract area functions obtained from the same speaker in 1994 and 2002. *J Acoust Soc Am*. 2008;123:327–335. https://doi.org/10.1121/1.2805683.

7. Story BH, Titze IR, Hoffman EA. Vocal tract area functions for an adult female speaker based on volumetric imaging. *J Acoust Soc Am*. 1998;104:471–487. https://doi.org/10.1121/1.423298.
8. Guzman M, Laukkanen AM, Krupa P, et al. Vocal tract and glottal function during and after vocal exercising with resonance tube and straw. *J Voice*. 2013;27:523.e19–523.e34. https://doi.org/10.1016/j.jvoice.2013.02.007.
9. Kitamura T, Takemoto H, Honda K. Acoustic characteristics of solid vocal tracts modeled from ATR MRI database of Japanese vowel production. *J Acoust Soc Am*. 2008;123:3734. https://doi.org/10.1121/1.2935241.
10. Takemoto H, Mokhtari P, Kitamura T. Acoustic analysis of the vocal tract during vowel production by finite-difference time-domain method. *J Acoust Soc Am*. 2008;123:3323. https://doi.org/10.1121/1.2933804.
11. Honda K, Kitamura T, Takemoto H. et al. Visualisation of hypopharyngeal cavities and vocal-tract acoustic modelling. *Comput Meth Biomech Biomed Eng*. 2010;13:443–453. https://doi.org/10.1080/10255842.2010.490528.
12. Hirai H, Honda K, Fujimoto I, et al. An MRI study of the physiological mechanism of *F*0 lowering. *J Acoust Soc Am*. 1993;93:2294–2295. https://doi.org/10.1121/1.406495.
13. Probst J, Lodermeyer A, Fattoum S. et al. Acoustic and aerodynamic coupling during phonation in MRI-based vocal tract replicas. *Appl Sci*. 2019;9:3562. https://doi.org/10.3390/app9173562.
14. Birkholz P, Kürbis S, Stone S, et al. Printable 3D vocal tract shapes from MRI data and their acoustic and aerodynamic properties. *Sci Data*. 2020;7:255. https://doi.org/10.1038/s41597-020-00597-w.
15. Takemoto H, Honda K, Saitou T, et al. Changes in the vocal tract shape of sopranos at high pitch. *J Acoust Soc Am*. 2012;131. https://doi.org/10.1121/1.4708739. 3378–3378.
16. Echternach M, Burk F, Burdumy M, et al. Morphometric differences of vocal tract articulators in different loudness conditions in singing. Bolhuis JJ, ed. *PLoS One*. 2016;11: e0153792. https://doi.org/10.1371/journal.pone.0153792.
17. Echternach M, Traser L, Richter B. Vocal tract configurations in tenors' passaggio in different vowel conditions—A real-time magnetic resonance imaging study. *J Voice*. 2014;28:262.e1–262.e8. https://doi.org/10.1016/j.jvoice.2013.10.009.
18. Echternach M, Sundberg J, Arndt S, et al. Vocal tract in female registers—A dynamic real-time MRI study. *J Voice*. 2010;24:133–139. https://doi.org/10.1016/j.jvoice.2008.06.004.
19. Echternach M, Sundberg J, Baumann T, et al. Vocal tract area functions and formant frequencies in opera tenors' modal and falsetto registers. *J Acoust Soc Am*. 2011;129:3955–3963. https://doi.org/10.1121/1.3589249.
20. Köberlein M, Birkholz P, Burdumy M. et al. Investigation of resonance strategies of high pitch singing sopranos using dynamic three-dimensional magnetic resonance imaging. *J Acoust Soc Am*. 2021;150:4191–4202. https://doi.org/10.1121/10.0008903.
21. Feng M, Howard DM. The dynamic effect of the valleculae on singing voice – An exploratory study using 3D printed vocal tracts. *J Voice*. 2021. https://doi.org/10.1016/j.jvoice.2020.12.012. Published online January.
22. Delvaux B, Howard D. A new method to explore the spectral impact of the Piriform Fossae on the singing voice: benchmarking using MRI-based 3D-printed vocal tracts. Robin DA, ed. *PLoS One*. 2014;9:e102680. https://doi.org/10.1371/journal.pone.0102680.
23. Seeing Voice: Visualization of vocal tract shaping with MRI (Meyer, Howard, Lingala – The voice foundation, *Care of the Professional Voice Symposium*, 2019. www.youtube.com. Accessed June, 2022. https://youtu.be/XlSVq7aQj2s.
24. Howard DM. The vocal tract organ and the vox humana organ stop. *J Music Technol Educ*. 2014;7:265–277. https://doi.org/10.1386/jmte.7.3.265_1.
25. Howard DM. Singing synthesis and the Vocal Tract Organ. eprints.whiterose.ac.uk. Published 2014. Accessed June, 2022. https://eprints.whiterose.ac.uk/87529/.
26. Howard, DM. Developing the vocal tract organ. *J Voice.*, doi: https://doi.org/10.1016/j.jvoice.2022.06.025 Accessed August, 2022.
27. Traser L, Birkholz P, Flügge TV, et al. Relevance of the implementation of teeth in three-dimensional vocal tract models. *J Speech Lang Hear Res*. 2017;60:2379–2393. https://doi.org/10.1044/2017_jslhr-s-16-0395.
28. Bresch E, Kim Yoon-Chul, Nayak K, et al. Seeing speech: capturing vocal tract shaping using real-time magnetic resonance imaging [Exploratory DSP]. *IEEE Signal Process Mag*. 2008;25:123–132. https://doi.org/10.1109/msp.2008.918034.
29. Isaieva K, Laprie Y, Leclère J, et al. Multimodal dataset of real-time 2D and static 3D MRI of healthy French speakers. *Sci Data*. 2021;8:258. https://doi.org/10.1038/s41597-021-01041-3.
30. Scott AD, Wylezinska M, Birch MJ, et al. Speech MRI: morphology and function. *Physica Med*. 2014;30:604–618. https://doi.org/10.1016/j.ejmp.2014.05.001.
31. Chen T, Chodara AM, Sprecher AJ. et al. A new method of reconstructing the human laryngeal architecture using micro-MRI. *J Voice*. 2012;26:555–562. https://doi.org/10.1016/j.jvoice.2011.03.012.
32. Regatte RR, Schweitzer ME. Ultra-high-field MRI of the musculoskeletal system at 7.0T. *J Magn Reson Imaging*. 2007;25:262–269. https://doi.org/10.1002/jmri.20814.
33. Lingala SG, Sutton BP, Miquel ME, et al. Recommendations for real-time speech MRI. *J Magn Reson Imaging*. 2015;43:28–44. https://doi.org/10.1002/jmri.24997.
34. Lingala SG, Toutios A, Töger J, et al. State-of-the-Art MRI protocol for comprehensive assessment of vocal tract structure and function. *Interspeech*. 2016. https://doi.org/10.21437/interspeech.2016-559. 2016. Published online September 8.
35. Burdumy M, Traser L, Burk F, et al. One-second MRI of a three-dimensional vocal tract to measure dynamic articulator modifications. *J Magn Reson Imaging*. 2016;46:94–101. https://doi.org/10.1002/jmri.25561.
36. Storck C, Gugatschka M, Friedrich G, et al. Developing a 3D model of the laryngeal cartilages using HRCT data and MIMICS's segmentation software. *Logoped Phoniatr Vocol*. 2010;35:19–23. https://doi.org/10.3109/14015430903552378.
37. Loth A, Corny J, Santini L, et al. Analysis of hyoid–larynx complex using 3D geometric morphometrics. *Dysphagia*. 2015;30:357–364. https://doi.org/10.1007/s00455-015-9609-2.
38. Vorik A, Unteregger F, Zwicky S, et al. Three-dimensional imaging of high-resolution computer tomography of singers' larynges—A pilot study. *J Voice*. 2017;31:115.e17–115.e21. https://doi.org/10.1016/j.jvoice.2016.03.011.
39. Unteregger F, Honegger F, Potthast S, et al. 3D analysis of the movements of the laryngeal cartilages during singing. *Laryngoscope*. 2016;127:1639–1643. https://doi.org/10.1002/lary.26430.
40. Newell JD, Fuld MK, Allmendinger T, et al. Very low-dose (0.15 mGy) chest CT protocols using the COPDGene 2 test object and a third-generation dual-source CT scanner with corresponding third-generation iterative reconstruction software. *Invest Radiol*. 2015;50:40–45. https://doi.org/10.1097/rli.0000000000000093.
41. Hammond E, Sloan C, Newell JD, et al. Comparison of low- and ultralow-dose computed tomography protocols for quantitative lung and airway assessment. *Med Phys*. 2017;44:4747–4757. https://doi.org/10.1002/mp.12436.
42. Hammond E, Chan KS, Ames JC, et al. Impact of advanced detector technology and iterative reconstruction on low-dose quantitative assessment of lung computed tomography density in a biological lung model. *Med Phys*. 2018;45:3657–3670. https://doi.org/10.1002/mp.13057.
43. Hoffman EA. Origins of and lessons from quantitative functional X-ray computed tomography of the lung. *Br J Radiol*. 2022;95. https://doi.org/10.1259/bjr.20211364.
44. Gordic S, Morsbach F, Schmidt B, et al. Ultralow-Dose chest computed tomography for pulmonary nodule detection: first performance evaluation of single energy scanning with spectral shaping. *Invest Radiol*. 2014;49:465–473. https://doi.org/10.1097/RLI.0000000000000037.
45. Radiology (ACR) RS of NA (RSNA) and AC of. Radiation dose in X-Ray and CT exams. *Radiologyinfo.org*. Published 2021. https://www.radiologyinfo.org/en/info/safety-xray.

46. Radioactivity: Radiation in Flight. www.radioactivity.eu.com. https://www.radioactivity.eu.com/site/pages/Radioactivity_in_Flight.htm.
47. Zhou X, Woo J, Stone M, et al. Improved vocal tract reconstruction and modeling using an image super-resolution technique. Published online 2013. doi:10.1121/1.4802903].
48. Arnela M, Guasch O, Alías F. Effects of head geometry simplifications on acoustic radiation of vowel sounds based on time-domain finite-element simulations. *J Acoust Soc Am*. 2013;134:2946–2954. https://doi.org/10.1121/1.4818756.
49. Blandin R, Arnela M, Félix S, et al. Efficient 3D acoustic simulation of the vocal tract by combining the multimodal method and finite elements. *IEEE Access*. 2022;10:69922–69938. https://doi.org/10.1109/ACCESS.2022.3187424.
50. Fleischer M, Mainka A, Kürbis S, et al. How to precisely measure the volume velocity transfer function of physical vocal tract models by external excitation. Reby D, ed. *PLoS One*. 2018;13: e0193708. https://doi.org/10.1371/journal.pone.0193708.
51. Jing B, Ge Z, Wu L, et al. Visualizing the mechanical wave of vocal fold tissue during phonation using electroglottogram-triggered ultrasonography. *J Acoust Soc Am*. 2018;143:EL425–EL429. https://doi.org/10.1121/1.5039726.
52. Tsai CG, Chen JH, Shau YW, et al. Dynamic B-Mode ultrasound imaging of vocal fold vibration during phonation. *Ultrasound Med Biol*. 2009;35:1812–1818. https://doi.org/10.1016/j.ultrasmedbio.2009.06.002.
53. Ye Z, Dun A, Jiang H, et al. The role of 3D printed models in the teaching of human anatomy: a systematic review and meta-analysis. *BMC Med Educ*. 2020;20:335.
54. Arai T. "Education in acoustics and speech science using vocal-tract models,". *J Acoust Soc Am*. 2012;131:2444–2454.
55. Arai T. "Vocal-tract models and their applications in education for intuitive understanding of speech production,". *Acoust Sci & Tech*. 2016;37:148–156.
56. Poblete P, McAleer S, Mason AG. 3D technology development and dental education: what topics are best suited for 3D learning resources? *Dent J*. 2020;8:95. https://doi.org/10.3390/dj8030095.
57. Howard DM. Raising public awareness of acoustic principles using voice and speech production. *J Acoust Soc Am – Invited Paper for Spec Issue on Acoust Educ*. 2012;131:2405–2412.