Tuning MRI-based vocal tracts to modify formants in the three-dimensional finite element production of vowels

Marc Arnela*, Oriol Guasch, Arnau Pont

GTM Grup de recerca en Tecnologies Mèdia, La Salle, Universitat Ramon Llull, Barcelona, Catalonia, Spain

Keywords: acoustic sensitivity functions; finite element method; MRI-based vocal tracts; vowels

Introduction

Accurate three-dimensional (3D) vocal tract (VT) geometries from Magnetic Resonance Imaging (MRI) (see e.g., [1]) can be used for the numerical simulation of sounds like vowels [2], diphthongs [3] and sibilants [4]. These VT geometries may need to be modified to match the formant frequencies from natural recordings, or to simply generate new sounds such as in singing voice. Vocal tract variations can be driven following the tuning strategy in [5], in which acoustic sensitivity functions were applied to 1D VTs described by an area function. In this work we extend the methodology in [5] to modify 3D MRI-based VTs by first converting them to 1D, then proceed to tuning and finally reverting back to 3D geometries.

Methods

The MRI-based VT geometries for static vowels sounds in [1] were selected for this work. First, they were adapted for our purposes, removing the face, lips, and subglottal tract. Second, they were discretized in a set of cross-sections along the vocal tract midline, as typically done to extract 1D area functions. However, not only the area of each cross-section was computed, but also their shape and location through the vocal tract midline. On the other hand, the mixed wave equation characterizing the 3D vowel generation problem was converted into a 1D Webster equation in the frequency domain, with complex wavenumber. Its eigenvalues and eigenvectors were obtained by means of the finite element method (FEM). A modal perturbation analysis followed which resulted in the same sensitivity functions of [5], but without the need to resort to non-linear radiation pressure. The sensitivity functions were then used in the iterative algorithm of [5], which induces variations in the 1D vocal tract area functions until formants reach prescribed target values. A 3D vocal tract geometry was then reconstructed from the final 1D area function, making use of the stored original shape and location of each VT cross-section.

FEM simulations were performed to analyze the vocal tract acoustics of the original and modified vocal tract geometries, using the in-house code outlined in [3]. The vocal tract transfer function $H(f)$ was computed for each configuration. Here, $H(f) = P_o(f)/Q_i(f)$, with $P_o(f)$ and $Q_i(f)$ respectively standing for the Fourier transform of the acoustic pressure at the mouth and the input volume velocity at the glottis.

Results

Figure 1 shows some results in which the MRI-based vocal tract of vowel [ɑ] is modified to produce a cluster of the third, fourth and fifth formants, as typically observed in singing voice.

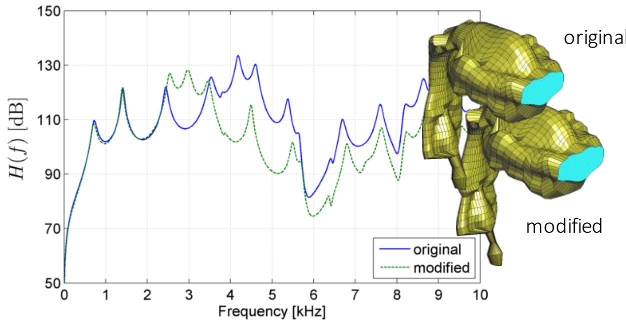

Figure 1: Vocal tract transfer function $H(f)$ of vowel [ɑ] for the original and modified 3D vocal tract geometries.

Conclusions

In this work, we have presented a methodology to tune the formants of 3D MRI-based vocal tracts, which has been applied to the FEM generation of vowel sounds. The swapping between 3D and 1D allows one to modify planar mode resonances, while keeping the higher frequency content of voice dictated by the 3D VT geometry.

Acknowledgements

This research has been supported by the Agencia Estatal de Investigación (AEI) and FEDER, EU, through project GENIOVOX TEC2016-81107-P, and by URL project 2019-URL-Proj-067.

References

[1] D. Aalto *et al*, "Large scale data acquisition of simultaneous MRI and speech," Appl. Acoust., 83:64–75, 2014.
[2] M. Arnela *et al*, "Influence of vocal tract geometry simplifications on the numerical simulation of vowel sounds," J. Acoust. Soc. Am., 140(3):1707–1718, 2016.
[3] M. Arnela *et al*, "MRI-based vocal tract representations for the three-dimensional finite element synthesis of diphthongs", IEEE/ACM Trans. Audio Speech. Lang. Process., 27(12):2173-2182, 2019.
[4] A. Pont *et al*, "Computational aeroacoustics to identify sound sources in the generation of sibilant /s/". Int. J. Numer. Meth. Biomed. Eng., 35(1):e3153, 2019.
[5] B. H. Story, Technique for "tuning" vocal tract area functions based on acoustic sensitivity functions, J. Acoustic. Soc. Am., 119(2):715-718, 2006.

*marc.arnela@salle.url.edu