3D Vocal tract reconstruction using magnetic resonance imaging data to study fricative consonants production

Sandra R. Ventura[1], Diamantino R. Freitas[2], Isabel M. Ramos[3], João Manuel R. S. Tavares[4]

[1] *Área Técnico-Científica da Radiologia, Escola Superior de Tecnologia da Saúde – Instituto Politécnico do Porto, sandra.rua@eu.ipp.pt*

[2] *Departamento de Engenharia Electrotécnica e de Computadores, Faculdade de Engenharia, Universidade do Porto, dfreitas@fe.up.pt*

[3] *Departamento de Radiologia do Centro Hospitalar de S. João EPE, Faculdade de Medicina, Universidade do Porto, radiologia.hsj@mail.telepac.pt*

[4] *Instituto de Engenharia Mecânica e Gestão Industrial, Departamento de Engenharia Mecânica, Faculdade de Engenharia, Universidade do Porto, tavares@fe.up.pt*

Introduction

The task of vocal tract modeling for articulatory synthesis systems intends to generate the complete geometrical information concerning the vocal tract measured by different technique, especially using Magnetic Resonance Imaging (MRI) [1, 2].
Currently, the use and application of MRI in speech research provide useful and accurate qualitative and quantitative analysis of speech articulation [3-6].
For a very long time, 3D modeling of vocal tract and speech production organs has been essentially limited to the midsagittal plane [7, 8]. However, with the improvement of MRI acquisition techniques, the three-dimensional (3D) modeling has become possible [9-11].
To obtain more detailed information between vocal tract shape and speech sounds, 3D data are required [12]. For this purpose, only the MRI technique provides excellent structural differentiation of all the vocal tract organs and without harmful effects.
The aim of this study is to describe an effective and useful method to extract the vocal tract volumes from MRI images during speech production. In addition, the reconstruction 3D models of the vocal tract are presented and analysed.

Material and methods

The MRI data used in this study was acquired at the Radiology Department of the Centro Hospitalar de S. João, Porto, Portugal, using a clinical 3.0 Tesla MRI system and a head and neck array coils.
Experiments were performed with two young volunteers (one male and one female) during some speech production tasks and according to the specific safety procedures.
Concerning the image processing and analysis tasks, the segmentation of the vocal tract was automatically obtained in each slice followed by the 3D reconstruction using the ITK-SNAP software.

Results and Discussion

The approach used to visualize the vocal tract during the production of fricative consonants in MR images, and the posterior 3D model reconstructed revealed to be very effective.
Quantitative measurements of volumes between each pair of fricative consonants (voiceless versus voiced) for each subject were also attained.
The approach used allows the visualization of the shape and volume of the vocal tract during the production of sustained consonants through 3D image-based reconstruction.
Comparing the 3D models of the vocal tract, it was found that the volumes measured for the voiceless consonants are smaller when compared to its counterpart voiced consonant. The opposite was observed, only for the male subject, during the production of the pairs of fricatives: post-alveolar and alveolar. Equivalent results were reported in [9] with the observation of some supraglottal volume changes of pharyngeal articulation during the production of voiced and voiceless fricative consonants.
Through the present development, volumetric MRI can be successfully applied to obtain 3D models, and measurable data concerning the vocal tract resonance cavities and all the articulators involved.

Due to the developments that have been occurred in the MRI systems, namely by the use of 3.0 Tesla magnetic fields, new applications and image refinements are expected, and consequently, significant improvements on the quality of the data acquired about the articulatory events during speech production.
This 3D knowledge of the speech organs could be very important, especially for clinical purposes (for example, for the assessment of articulatory impairments followed by tongue surgery in speech rehabilitation), and also for a better understanding of the acoustic theory in speech production.

References

[1] B. Kröger and P. Birkholz, "Articulatory synthesis of speech and singing: State of the art and suggestions for future research," *Multimodal Signals: Cognitive and Algorithmic Issues*, pp. 306–319, 2009.

[2] S. S. Narayanan, A. A. Alwan, and K. Haker, "An articulatory study of fricative consonants using magnetic resonance imaging," *Journal of the Acoustical Society of America*, vol. 98, no. 3, pp. 1325–1347, 1995.

[3] S. R. Ventura, D. R. Freitas, I. M. Ramos, and J. M. R. S. Tavares, "Morphological Differences in the Vocal Tract Resonance Cavities of Voice Professionals: An MRI- based Study," *Journal of Voice*, vol. 27, no. 2, pp. 132–140, 2013.

[4] S. M. R. Ventura, D. R. S. Freitas, and J. M. R. S. Tavares, "Toward dynamic magnetic resonance imaging of the vocal tract during speech production," *Journal of voice*, vol. 25, no. 4, pp. 511–518, 2011.

[5] S. Ventura, D. Freitas, M. Ramos, and J. M. R. S. Tavares, "Requisitos e condicionantes da imagem por ressonância magnética no estudo da fala humana," in *Congresso de Métodos Numéricos em Engenharia (CMNE)*, 2011, pp. 1–12.

[6] M. J. Vasconcelos, S. M. Ventura, D. R. Freitas, and J. M. R. Tavares, "Inter-speaker speech variability assessment using statistical deformable models from 3.0 Tesla magnetic resonance images," *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, vol. 226, no. 3, pp. 185–196, Dec. 2011.

[7] A. Serrurier, P. Badin, and others, "Towards a 3D articulatory model of velum based on MRI and CT images," *Papers in Linguistics*, vol. 40, no. 1, pp. 195–211, 2005.

[8] P. Badin, G. Bailly, M. Raybaudi, and C. Segebarth, "A three-dimensional linear articulatory model based on MRI data," in *Proceed. 5th International Conference on Spoken Language Processing (ICSLP 98)*, 1998, vol. 2, pp. 417–420.

[9] M. I. Proctor, C. H. Shadle, and K. Iskarous, "Pharyngeal articulation in the production of voiced and voiceless fricatives," *The Journal of the Acoustical Society of America*, vol. 127, no. 3, pp. 1507–18, Mar. 2010.

[10] P. Badin, A. Serrurier, and others, "Three-dimensional modeling of speech organs: Articulatory data and models," *IEIC Technical Report (Institute of Electronics, Information and Communication Engineers)*, vol. 106, no. 177, pp. 29–34, 2006.

[11] K. Van Den Doel, F. Vogt, and R. English, "Towards Articulatory Speech Synthesis with a Dynamic 3D Finite Element Tongue Model," *Computer*, pp. 59–66, 2006.

[12] S. R. Ventura, D. R. Freitas, and J. M. R. S. Tavares, "Application of MRI and biomedical engineering in speech production study.," *Computer methods in biomechanics and biomedical engineering*, vol. 12, no. 6, pp. 671–81, Dec. 2009.