Improved Depiction of Tissue Boundaries in Vocal Tract Real-time MRI Using Automatic Off-resonance Correction

Yongwan Lim, Sajan Goud Lingala, Asterios Toutios, Shrikanth Narayanan, and Krishna Nayak

Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, CA, USA

Abstract

- Real-time magnetic resonance imaging (RT-MRI) is a powerful tool to study the dynamics of vocal tract shaping during speech production.
- The dynamic articulators of interest located at air-tissue interfaces include the surfaces of the lips, tongue, hard palate, soft palate, and pharyngeal airway and are vulnerable to MRI ***off-resonance effect*** due to magnetic susceptibility.
- In RT-MRI using spiral scanning, ***off-resonance effect*** appears as a signal loss or blurring in images and may impair the analysis of dynamic speech data.
- We apply an automatic off-resonance artifact correction method to speech RT-MRI data in order to enhance the sharpness of air-tissue boundaries.
- We demonstrate an improvement of depiction of the vocal tract area in alveolar ridge, hard-, soft-palate, and the tongue boundaries, qualitatively and using an image sharpness metric.

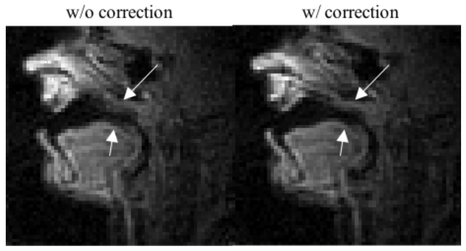

Fig. 1. An example frame of RT-MRI images without and with the off-resonance correction.

Motivation

- **Off-resonance artifact has a significant potential impact on**
  1. RT-MRI analysis in speech science
     - Air-tissue boundary segmentation [1]
     - Biomarkers such as average pixel intensity [2]
     - Assessment of velopharyngeal insufficiency
  2. Current speech RT-MRI [3,4]
     - Lower field strength (1.5 Tesla) MRI scanner
     - Short duration (~2.5ms) readouts [3] → multi-shot spirals → limited scan efficiency
- **Off-resonance correction may improve**
  - **1) the analysis of articulator dynamics, 2) imaging efficiency in RT-MRI**

Materials and Methods

- **USC-TIMIT datasets** [5]
  - Previously collected RT-MRI data of the vocal tract during speech production
  - 3 speakers : the most significant blurring artifacts on visual assessment
  - Imaging parameters
    - A 13-interleaf spiral sequence in the mid-sagittal plane
    - TR/TE = 6.164 /0.8 ms, SR = 2.4mm^2, FOV = 200mm^2, Tro = 2.52ms
    - A sliding window technique with a frame rate of 23.18 frames/s.

- **Automatic Off-resonance Correction**
  - *Off-resonance effect*
    - Deviation in resonance frequency from the signal receive frequency in MRI due to the magnetic susceptibility difference between air and tissue
  - Automatic off-resonance correction (Auto-focus)
    - Reconstruct artifact-corrected image without a prior knowledge about local off-resonance frequency (field map) [6,7]
    - Field map estimation performed using [7] with a modified focus metric:

$$\text{Focus metric}: S(x,y,t;\Delta f_i) = \sum_{(x,y,t)\in A(x,y,t)} |Im\{I(x,y,t;\Delta f_i)\}|$$

$$\text{Field map}: \Delta f(x,y,t) = \underset{\Delta f_i}{\operatorname{argmin}} S(x,y,t;\Delta f_i)$$

$$\text{Corrected image}: \bar{I}(x,y,t) = I\big(x,y,t;\Delta f(x,y,t)\big)$$

    - $\{\Delta f_i, i = 1,2,\ldots,N_f\}$ is a set of equally spaced frequencies
    - $A(x, y, t)$ is a $w_x \times w_y \times w_t$ summation window centered at (x,y,t),
    - $I(x,y,t;\Delta f_i)$ is the image reconstructed at a frequency Δf_i
    - Parameters used
      - w_x (= w_y) = 21~25, w_t = 7~11 for coarse estimation
      - w_x (= w_y) = 7~9, w_t = 7~11 for fine estimation
      - The range of Δf_i: [−120 120] Hz, N_f = 13

- **Sharpness Evaluation**
  - Gridlines (Fig. 2(a)) extracted at air-tissue boundaries using the grid extraction method [1]
  - Edge sharpness at air-tissue boundaries
    - The reciprocal of the distance (1/d) between the points of 80% and 20% of the maximum intensity value (Fig. 2(b)) and averaged over all time frames

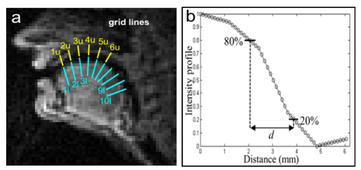

Fig. 2. Sharpness measure. (a) Extraction of gridlines of upper and lower boundaries. (b) Intensity profile of grid line

Results – Qualitative Assessment

- Posterior to the alveolar ridge, up to the soft palate becomes more intense and sharper compared to the original images.
- The tongue boundary appears sharper in the corrected images.

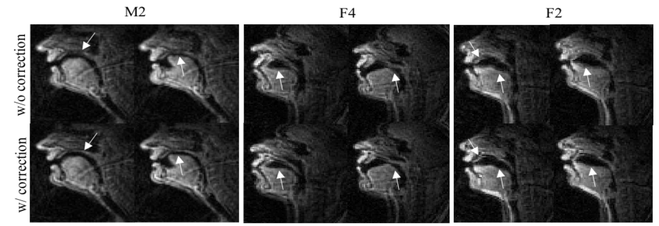

Fig. 3. Representative mid-sagittal image frames of vocal tracts for three speakers

- The intensity profile exhibits
  - 1) sharper the tongue boundary, 2) a clear delineation of the soft palate movements
- For speaker M2, the intensity in the hard palate shows more constant along time
  - This result agrees with the fact that the hard palate, which is a bony structure covered by a thin layer of tissue, does not change its shape during speech production [8].

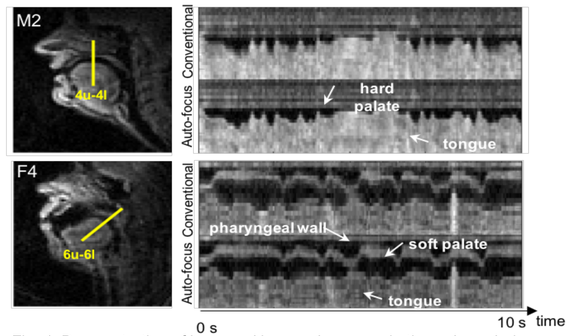

Fig. 4. Demonstration of improved image sharpness in tissue boundaries in intensity profile along time

Results – Quantitative Assessment

- Upper boundaries
  - Higher sharpness scores at the hard and soft palate
- Lower boundaries
  - Significant improvements at the middle and back tongue boundaries
  - Less or lower sharpness at the front tongue boundaries

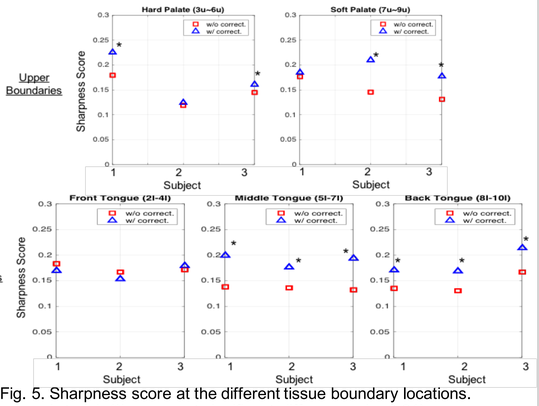

Fig. 5. Sharpness score at the different tissue boundary locations. Asterisk (*) represent significant differences (p<0.01)

Conclusion

- An automatic off-resonance correction method shows improved image depiction of the vocal tract area such as the tongue boundaries, alveolar ridge, and soft palate, which has the potential to improve the analysis of the dynamics of articulators.

Acknowledgments

- This work was supported by NIH grant R01DC007124 and NSF grant 1514544.
- We acknowledge the support and collaboration of the Speech Production and Articulation kNowledge (SPAN) group at University of Southern California, Los Angeles, CA, USA.

References

[1] J. Kim et al, *ISSP*, pp. 222–225, 2014.
[2] M. Proctor et al, *Interspeech*, pp. 281–284, 2011.
[3] S. Lingala et al, *Magn. Reson. Med.*, doi: 10.1002/mrm.26090, 2016.
[4] S. Narayanan et al, *JASA*, 115(4): 1771–1776, 2004.
[5] S. Narayanan et al, *JASA*, 136(3): 1307–1311, 2014.
[6] D. C. Noll et al, *Magn. Reson. Med.*, 25(2):319–333, 1992.
[7] L. C. Man et al, *Magn. Reson. Med.*, 37(6):906–913, 1997.
[8] E. Bresch et al, *IEEE Trans. Med. Imaging*, 28(3):323–338, 2009.