Time-interleaved imaging of arbitrary scan planes applied to real-time speech MRI

Y-C. Kim[1], M. I. Proctor[1], S. S. Narayanan[1], and K. S. Nayak[1]
[1]Department of Electrical Engineering, University of Southern California, Los Angeles, CA, United States

Introduction Current methods for real-time MRI of the upper airway for speech research involve imaging a single (typically midsagittal) slice [1-4], allowing for extraction of parameters related to vocal tract shaping. This provides insights into the dynamics of all articulators, but does not allow for visualization of several important features in vocal tract shaping such as grooving/doming of the tongue, asymmetries in tongue shape, and lateral shaping of the pharyngeal wall. In this work, we image multiple scan planes in time-interleaved fashion, compromising temporal resolution for greater spatial coverage. We demonstrate several advantages of this approach to the study of vocal tract dynamics.

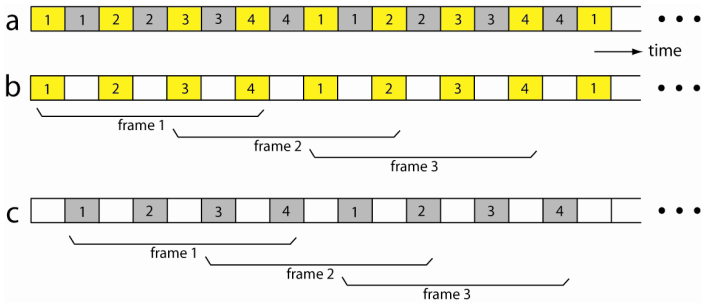

Figure 1. *Illustration of the slice-interleaf order. Slice 1 and 2 are denoted by yellow and gray colors, respectively. The spiral interleaf index is indicated by the number within each block. This shows 4-interleaf case, and 13-interleaf spiral was used in the in-vivo experiments. (a) Temporal ordering of the slice and spiral interleaf. Sliding window reconstruction for (b) slice 1 and (c) slice 2. Frames 1, 2, and 3 in (b) and (c) are simultaneous, for all practical purposes.*

Methods Experiments were performed on a GE Signa Excite 1.5 T scanner with a custom 4-channel receiver coil (two anterior elements, two posterior elements) using custom real-time imaging software [5]. Imaging parameters were: FOV = 20 × 20 cm^2, flip angle = 15°, slice thickness = 6 mm, spiral gradient echo sequence, 13-interleaf uniform density spiral trajectories, 2.4 × 2.4 mm^2 spatial resolution, TR = 6.004 ms, sequential view order. Slice-interleaf order followed the scheme shown in Fig. 1. Temporal resolution was 6.004 * 26 = 156.1 ms. Real-time video was obtained at a frame rate of 23.8 fps after a sliding window reconstruction.

Imaging data were acquired from two subjects: one native speaker of Mandarin Chinese and one native English speaker. Stimuli were presented in-scanner, using a mirror-projector setup, and uttered at a normal speech rate. Two slices of interest were prescribed and saved during real-time localization, and then used for the time-interleaved imaging. Synchronized audio recordings were collected and gradient noise was removed during post-acquisition processing [6]. Mandarin sibilant fricatives /s/, /ʂ/, /ɕ/ were elicited in symmetrical maximally-contrastive vocalic contexts /a_a/, /i_i/, and /u_u/ [7], using pinyin stimuli. English fricative pairs /f-v/, /θ- ð /, /s-z/, /ʃ- ʒ / were elicited in symmetrical maximally-contrastive vocalic contexts /a_a/, /i_i/, /u_u/ using pseudo-word stimuli.

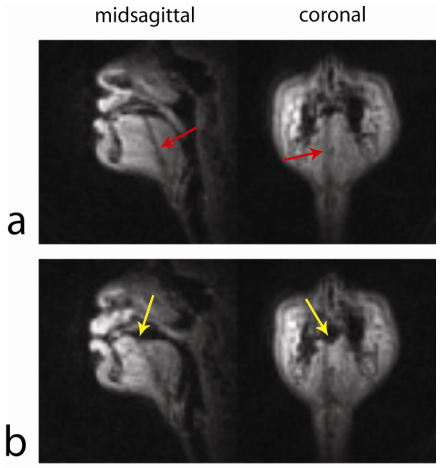

Figure 2. *Frames captured when the subject was (a) stationary and (b) producing the speech sound /s/.*

Results Figure 2 illustrates partial saturation of spins (see red arrows), which provides useful indication of the relative position of the slices. For example, both coronal images show that the prescribed midsagittal slice is slightly off the mid-line. Tongue grooving is not seen in the midsagittal slice, but is clearly seen in the coronal slice (see yellow arrows).

Vocal tract shaping during production of voiceless /s/ and voiced /z/ fricative pairs is compared in Figure 3. The constriction location at the alveolar ridge is almost the same for /s/ and /z/ (see yellow arrows in (a) and (c)). A larger pharyngeal area and more pronounced grooving (white arrows) of the tongue root is observed in the voiced sibilant fricative /z/ (compare (b) and (d)). Although not shown in this abstract, in the axial slice more rounded shape of the pharyngeal wall was observed for the voiced /ʒ/ than for the voiceless /ʃ/.

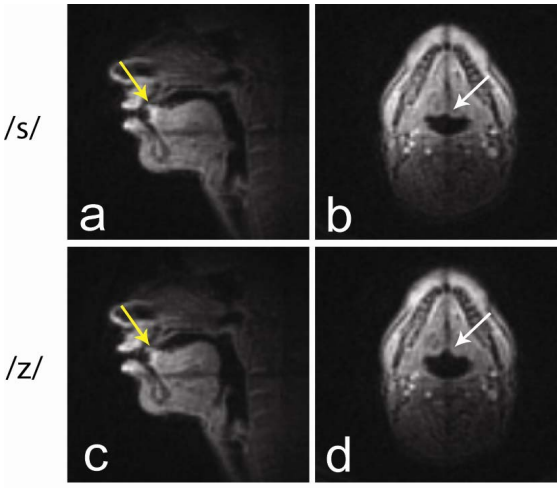

Figure 3. *Midsagittal and axial frames captured for /s/ and /z/ when producing /busubu/ and /buzubu/, respectively. The frames (a) and (b) were from the same time point. So were the frames (c) and (d).*

Discussion Interleaved real-time MRI of multiple slices provides valuable information about lingual articulation and vocal tract shaping, compared to conventional single-slice midsagittal imaging. The technique allows for more detailed analysis of consonant production using more natural stimuli, without requiring subjects to sustain articulatory postures for artificially long durations, as is typically required in conventional static 3D MR imaging study [8,9]. Time-interleaved imaging can also be extended to more than two scan planes with further sacrifices in temporal resolution. Real-time adjustment of the scan planes (e.g., perfect midsagittal) from knowledge of location of the partially saturated spins can provide more accurate prescription of the scan planes, and its implementation remains as future work.

Acknowledgments This work is supported by NIH Grant R01 DC007124-01.

References [1] Narayanan et al., J. Acoust. Soc. Am. 115:1771-1776(2004); [2] Bresch et al., IEEE SPM 25:123-132(2008); [3] Sutton et al., JMRI 32:1228-1237(2010); [4] Uecker et al., NMR in Biomed 23:986-994(2010); [5] Santos et al., IEEE EMBS 2:1048-1051(2004); [6] Bresch et al., J. Acoust. Soc. Am. 120:1791-1794(2006); [7] Öhman, J. Acoust. Soc. Am. 39:151-168(1966); [8] Proctor et al., J. Acoust. Soc. Am. 127:1507-1518(2010); [9] Narayanan et al., J. Acout. Soc. Am. 98:1325-1347(1995).