

High spatio-temporal resolution multi-slice real time MRI of speech using golden angle spiral imaging with constrained reconstruction, parallel imaging, and a novel upper airway coil

Sajan Goud Lingala¹, Yinghua Zhu¹, Yoon-Chul Kim², Asterios Toutios¹, Shrikanth Narayanan¹, and Krishna S Nayak¹

¹Electrical Engineering, University of Southern California, Los Angeles, California, United States, ²Samsung Medical Center, Seoul, Korea

Target audience: Researchers interested in rapid imaging of speech.

Purpose/Introduction: Real time MRI (RT-MRI) is a powerful tool to visualize the complex spatio-temporal coordination of upper airway structures during speech production. It is continuing to provide new insights in several linguistic studies, and also has potential to inform treatment plans in several clinical applications (eg. clefts of lips/palate, oropharyngeal cancer). RT-MRI is limited by intrinsic trade-offs in spatio-temporal resolution, signal to noise ratio, and slice coverage. Recently, sparse sampling and constrained recovery methods have shown promise to improve image quality in RT-MRI [1-2]. In this work, we utilize advances in coil design, non-Cartesian sampling, and compressed sensing to improve these tradeoffs. We employ (a) a novel upper airway coil which has improved sensitivity in upper airway regions of interest, (b) a flexible slice selective spiral acquisition with golden angle time interleaving, and (c) a temporal constrained reconstruction scheme. We demonstrate the proposed approach can dramatically improve the visualization of rapid articulatory movements, such as the production of consonant clusters. We also show the approach enables simultaneous imaging of the axial, mid-sagittal, and coronal slice planes, which in turn enables visualization of complex movements such as tongue grooving, and airway narrowing with high time resolution (up to 36ms).

Methods: Experiments were performed on a GE Signa Excite 1.5 T scanner with a custom 8-channel upper airway receiver coil that has four elements on either side of the jaw; the coil was designed to provide high sensitivity over all the important articulators (lips, tongue, epiglottis, velum). A gradient echo based spiral acquisition with golden angle ordering [3] was employed (FOV: 20cm², FA: 15 degrees, slice thickness: 5mm, TR=6.004ms). Single slice sequences with spatial resolutions of 2.5mm² and 1.34mm² were realized by modifying the number of spiral interleaves required to satisfy Nyquist (R=1) sampling. An interleaved three-slice sequence at 2.5mm² was also realized [4]. In Fig.1, we demonstrate the spatial and temporal resolution trade-offs with spiral sequences that are designed to make maximum use of gradients (40 mT/m amplitude and 150 mT/m/ms slew rate). As shown in Fig. 1, we propose to realize a reduction factor R~6.5. We utilized a sparse SENSE image reconstruction strategy with temporal total variation as the sparsity constraint [5]. The coil maps were estimated from time averaged data using an eigen-decomposition method [6], and the resulting optimization was solved by using a non-linear conjugate gradient algorithm. We chose the regularization parameter empirically. Image data were acquired from a healthy 29 year old adult male speaker during the following stimuli production (a) repetition of phrase: /loo-/lee-/laa-/za-/za (b) counting of numbers: one-two-three-four at normal pace, followed by a rapid pace.

Results: In Figures 2,3, we show the single slice reconstructions when the subject was counting numbers at different rates. We observe the proposed accelerated images to depict the rapidly changing temporal dynamics considerably better than the images formed at R=1. Note the superior temporal fidelity in terms of crispness in the time intensity profiles. In particular, the R=1 images shows considerable blurring at the tongue tip during the production of the word “three” (consonant cluster at the onset), while the accelerated images showed no apparent blurring, and improved fidelity. As depicted in Fig.3, there is a trade-off with the signal to noise level as spatial resolution is increased; the denoising effect of constrained reconstruction partly reduces this trade-off. In Fig. 4, with the utterance /loo-/lee-/laa-/za-/na-/za, notice the changes such as tongue grooving, airway narrowing to be well captured in the coronal, axial slices; these can provide valuable information in addition to the dynamics in a single mid-sagittal plane.

Conclusions: We have demonstrated a sparse spiral sampling and constrained reconstruction approach to improve the spatial, temporal resolutions and slice coverage in RT-MRI of speech. In comparison to Nyquist sampling, the proposed approach provides improved depiction of fine articulatory movements by enabling time resolution of 12 msec for single slice imaging, and 36 msec for concurrent three slice imaging. While the proposed approach shows tremendous potential, the reconstructed images need to be characterized in terms of residual artifacts, and blurring with an application driven hypothesis, and these remain as future work.

References: [1] M.Fu et al. MRM 2014 (early view, June 2014), [2] A. Niebergall et al, vol. 69(2):477-85, MRM 2013, [3] Y.Kim et al, vol.65(5): 1365-71, MRM 2011, [4] Y.Kim et al, vol.35(4): 943-8, JMIR 2012, [5] L.Feng et al, vol.72(3): 707-17, MRM 2014, [6] D.Walsh et al, vol.43(5) : 682-90, MRM 2000.

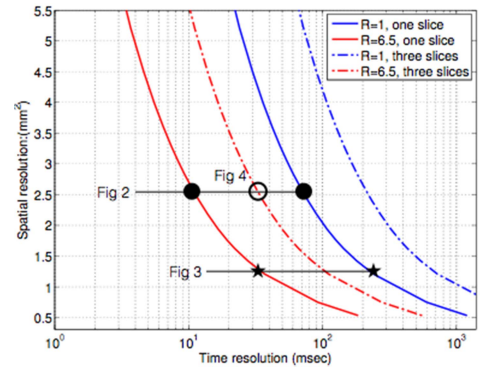


Fig.1: Sparse spiral sampling reduces spatial and temporal resolution trade-offs in RT-MRI of speech.

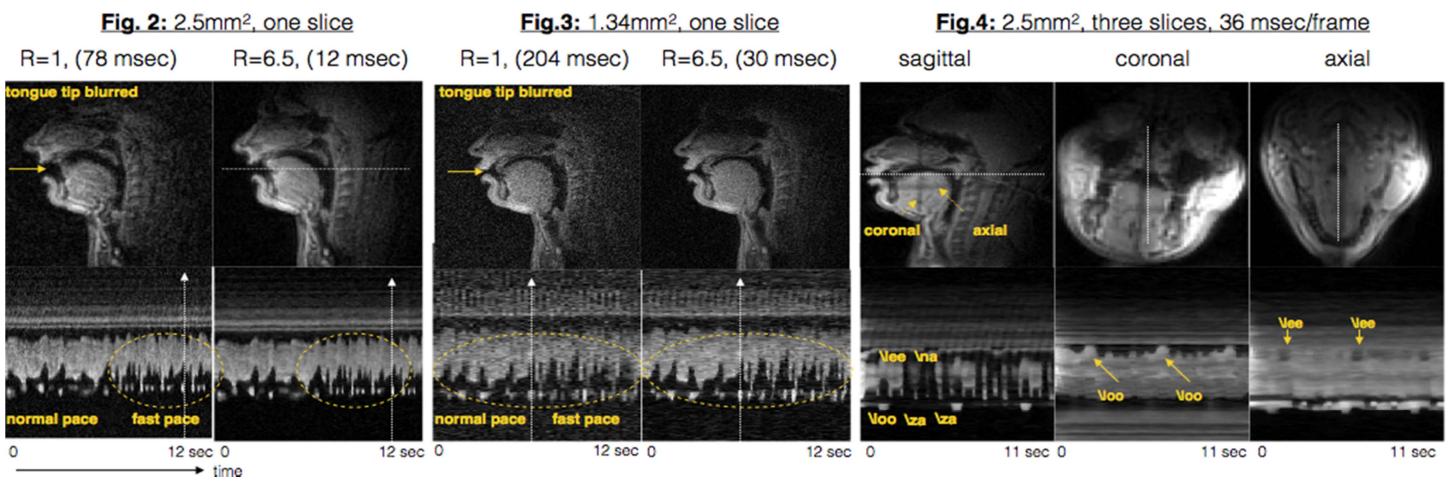


Fig.2,3: Image quality comparison of the conventional Nyquist reconstructed, and proposed constrained reconstruction based images: The top row shows spatial frames at instances depicted by the white arrows in the bottom row. The bottom row shows the image time profiles for the cross-section depicted in top row. The subject uttered “one-two-three-four” at a normal pace, followed by a rapid pace. Note the improvement in temporal fidelity in the time profile plots of the accelerated images (crispness along time in the yellow circles). Fig.4: Simultaneous visualization of sagittal, coronal, and axial planes during production of /loo-/lee-/laa-/za-/na-/za sounds with a time resolution of 36 msec. Note tongue grooving and airway narrowing can be visualized well in the coronal and axial planes.