Pseudo golden-ratio spiral imaging with gradient acoustic noise cancellation: application to real-time MRI of fluent speech

Yoon-Chul Kim[1], Nassos Katsamanis[1], Michael Proctor[1], Shrikanth Narayanan[1], and Krishna S. Nayak[1]
[1]*Electrical Engineering, University of Southern California, Los Angeles, CA, United States*

Introduction In the context of real-time MRI of the vocal tract during fluent speech, golden-ratio (GR) spiral imaging has been shown to provide improved depiction of articulators over conventional spiral imaging [1]. However, one major drawback of GR spiral imaging is incompatibility with MRI-gradient acoustic noise cancellation methods [2] because there is no periodicity in the gradient waveforms with respect to time. We apply GR spiral view ordering with a fixed number of interleaves (i.e. pseudo golden-ratio) to promote periodicity.

Methods Experiments were performed on a GE 1.5T scanner with a 4-channel upper airway coil using custom real-time imaging software [3,6]. The spiral trajectory design parameters were: 13-interleaf spiral, 20×20 cm^2 FOV, 2.4×2.4 mm^2 spatial resolution. Other parameters were: 6 mm slice thickness, gradient-echo sequence, 6.004 ms TR. The following acquisitions were used: (i) 13-interleaf spiral with pseudo-random temporal view order (Fig 1a) and (ii) pseudo-GR spiral trajectories with 21, 34, and 55 interleaves. We chose the Fibonacci numbers because the angle between the beginning and end of interleaves tends to be close to the golden-angle (≈ 222.49°) (Fig 1d). The angle more closely approximates the golden-angle for larger Fibonacci numbers.

A native English speaker was imaged while producing five utterances from the 460-sentence TIMIT corpus, presented using a mirror-projector setup [4]. Synchronized audio recordings were made during MRI scanning, and speech signals were obtained after noise cancellation [2]. Noise-cancelled audio tracks from the 4 pulse sequences were each phonetically transcribed at segment and word levels using forced-alignment software robust to degraded speech recordings [4,7]. Data from all 4 coil elements were used for image reconstruction. Sum-of-squares of gridding-reconstructed images was used for the 13-intl case. L1-SPIRiT reconstruction [5] was used for the pseudo GR cases. We retrospectively chose 78 ms and 48 ms resolutions. Coil calibration information was drawn at every temporal frame using a large temporal window. All videos were reconstructed at a frame rate of 55.5 fps using an 18 ms sliding window.

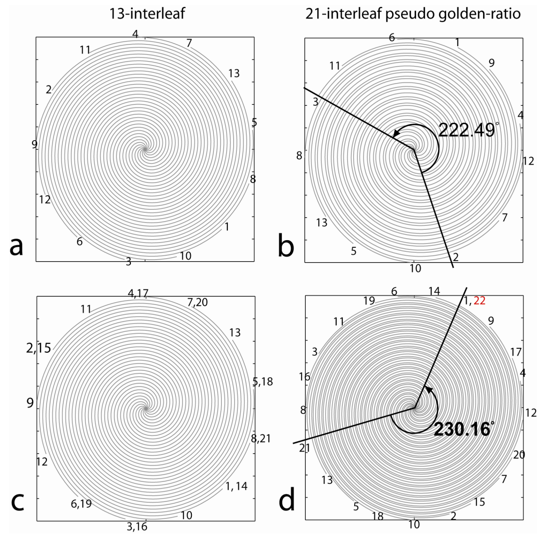

Fig 1. Comparison of (a,c) conventional 13-intl and (b,d) 21-intl pseudo GR spiral trajectory. Temporal view order is indicated by the number at the end of interleaf. Sampling patterns for (a,b) 13 and (c,d) 21 consecutive spiral interleaves. The spiral angle increment between 21 and 22 is not equal to the golden angle.

Results The quality of the noise-cancelled audio tracks degraded as the number of spiral interleaves increased. From the track of a 'silence' period (i.e. MRI scans when the subject was silent), relative noise power for the 21-, 34-, 55-interleaf with respect to the conventional case was 1.3dB, 6.9dB, 12.2dB high, respectively. L1-SPIRiT reconstruction was effective at removing aliasing artifacts that are observed in the conventional method (see yellow solid arrow in Fig 2a). Inspection of midsagittal articulatory configuration reveals some misalignment of audio and video frames. For example, see tongue tip release indicated by hollow arrow in Fig 2d for the aligned phoneme /dʒ/ which requires tongue tip constriction in the articulation. Fig 3 shows image frames captured in production of /n/ during the utterance of "money". Depiction of tongue tip motion was improved with the pseudo GR scheme when selecting a small temporal window for reconstruction (compare arrows in Fig 3).

Discussion The proposed method (for the case of 21 or 34 interleaf) provides improved image quality and adequate sound quality (and/or comparable alignment accuracy) compared to the conventional method. Although the current acoustic processing algorithm was not designed for non-periodic MRI sequences [6], the noise-cancelled audio signal obtained using the technique is of adequate quality to allow for the application of pseudo GR acquisition to a broader range of speech imaging studies, where the configurable temporal and spatial resolution of fast moving articulators is critical. The refinement of the pseudo GR spiral method to a greater range of speech studies remains as future work (for more information, see for example: http://sail.usc.edu/span/mri-timit/).

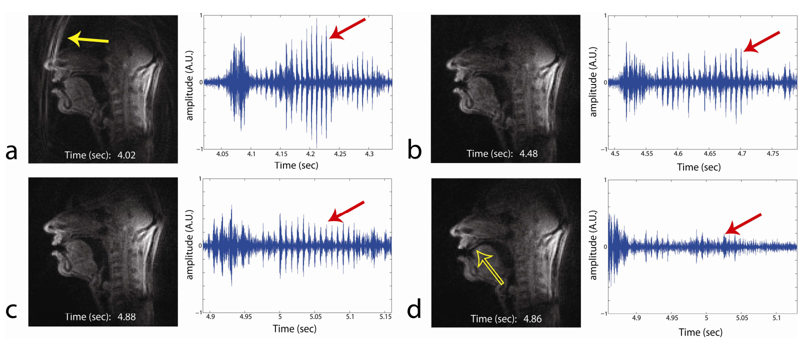

Fig 2. A frame corresponding to the onset in the alignment result of the word "Jane" and audio track corresponding to the word "Jane" from (a) the conventional 13 intl. spiral and pseudo GR spirals with (b) 21, (c) 34, (d) 55 intl. The periodic signals from voicing are seen in (a-c) but not seen in (d) (see red arrows).

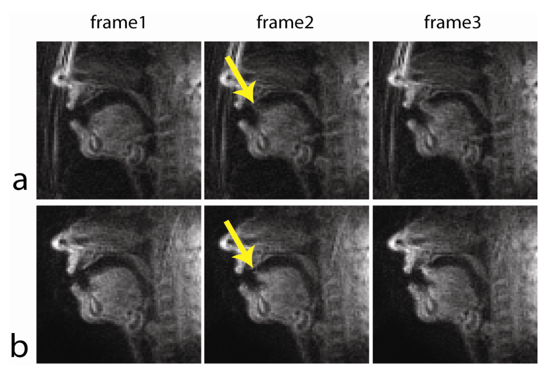

Fig 3. Comparison of tongue tip dynamics for (a) conventional 13 intl spiral method with 78 ms resolution and (b) pseudo GR 34 intl spiral method with the selection of 48 ms resolution.

Acknowledgments The work is supported by NIH Grant R01-DC007124.

References [1] Kim et al., *MRM* 65:1365-1371 (2011). [2] Bresch et al., *J Acoust Soc Am* 120:1791-1794 (2006). [3] Santos et al., *IEEE EMBS* 1048-1051 (2004). [4] Narayanan et al., *Interspeech* 12:837-840 (2011). [5] Lustig et al., *MRM* 64:457-471 (2010). [6] Bresch et al., *IEEE Sig. Proc. Mag.* 123-132 (2008). [7] Katsamanis et al., *Workshop on New Tools and Methods for VLSPR* (2011).