



Published in final edited form as:

Magn Reson Med. 2011 May ; 65(5): 1365–1371. doi:10.1002/mrm.22714.

Flexible Retrospective Selection of Temporal Resolution in Real-time Speech MRI Using a Golden-Ratio Spiral View Order

Yoon-Chul Kim*, Shrikanth S. Narayanan, and Krishna S. Nayak

Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089-2564

Abstract

In speech production research using real-time MRI, the analysis of articulatory dynamics is performed retrospectively. A flexible selection of temporal resolution is highly desirable because of natural variations in speech rate and variations in the speed of different articulators. The purpose of the study is to demonstrate a first application of golden-ratio spiral temporal view order to real-time speech MRI and investigate its performance by comparison with conventional bit-reversed temporal view order. Golden-ratio view order proved to be more effective at capturing the dynamics of rapid tongue tip motion. A method for automated blockwise selection of temporal resolution is presented that enables the synthesis of a single video from multiple temporal resolution videos and potentially facilitates subsequent vocal tract shape analysis.

Keywords

Golden ratio; vocal tract shaping; real-time spiral imaging; speech production

INTRODUCTION

Real-time MRI has provided new insight into the dynamics of vocal tract shaping during natural speech production (1–4). In real-time speech MRI experiments, image data and speech signals are simultaneously acquired. Real-time movies, typically of a 2D midsagittal slice, are reconstructed and displayed in real-time. The shape of the vocal tract, from the lips to the glottis, is identified using air-tissue boundary detection performed at each frame (5). Adaptive noise cancellation is used to produce speech signals free from the MRI gradient noise (6). Articulatory and acoustic analysis is then performed using synchronized audio and video information (7). Although MRI data is acquired and reconstructed in real-time, the processes of segmentation and analysis are performed retrospectively.

Speech rate is highly dependent on the subject's speaking style and the speech task, and it affects speed of articulatory movement (8,9). Variations in the velocity of articulators such as tongue dorsum, lips, and jaw result from the nature of the sequences of the vowels and consonants being produced (10). The motion of articulators (e.g., tongue, velum, lips) is relatively slow during production of monophthongal vowel sounds or during/vicinity of pauses. Vocal tract variables such as tongue tip constriction, lip aperture, and velum aperture are dynamically controlled and coordinated to produce target words (11). The speeds among

*Correspondence to: Yoon-Chul Kim, M.S., 3740 McClintock Ave, EEB 400, University of Southern California, Los Angeles, CA 90089-2564, Phone: (213) 740-4650, Fax: (213) 740-4651, yoonckim@usc.edu.

Preliminary portions of this work were presented at ISMRM 2010 abstract #4967.

articulators can also differ during the coordination of different articulators, for example, the movement of the velum and the tongue tip during the production of the nasal /n/.

Current speech MRI protocols do not provide a mechanism for flexible selection of temporal resolution. This is of potential value, because higher temporal resolution is necessary for frames that reflect rapid articulator motion while lower temporal resolution is sufficient for capturing the frames that correspond to static postures. As recently shown by Winkelmann et al. (12), golden-ratio sampling enables flexible retrospective selection of temporal resolution. It may be suited for speech imaging, in which the motion patterning of articulators varies significantly in time, and in which it is difficult to determine an appropriate temporal resolution *a priori*.

In this manuscript, we present a first application of spiral golden-ratio sampling scheme to real-time speech MRI and investigate its performance by comparison with conventional bit-reversed temporal view order sampling scheme. Simulation studies are performed to compare unaliased field-of-view (FOV) from spiral golden-ratio sampling with that from conventional bit-reversed sampling at different levels of temporal resolution after a retrospective selection. *In vivo* experiments are performed to qualitatively compare image signal-to-noise ratio (SNR), level of spatial aliasing, and degree of temporal fidelity. Finally, we present an automated technique in which a composite movie can be produced using data reconstructed at several different temporal resolutions. We demonstrate its effectiveness at improving articulator visualization during production of nasal consonant /n/.

MATERIALS and METHODS

Simulation

A simulation study was performed to compare unaliased FOVs from conventional bit-reversed view order sampling (5,13,14) (see Fig. 1a) and spiral golden-ratio view order sampling (see Fig. 1b) for a variety of temporal resolutions selected retrospectively. The spiral trajectory design was based on the imaging protocol routinely used in our laboratory at the University of Southern California (3,5,13). The design parameters were: 13-interleaf uniform density spiral (UDS), $20 \times 20 \text{ cm}^2$ FOV, $3.0 \times 3.0 \text{ mm}^2$ in-plane spatial resolution, maximum gradient amplitude = 22 mT/m, maximum slew rate = 77 T/m/s, and conventional bit-reversed view order. Bit-reversed temporal view order is often adopted in real-time MRI because it shortens the spiral interleaf angle gaps in a few adjacent interleaves at any time point and reduces motion artifacts (15,16). Spiral golden-ratio view order was performed by sequentially incrementing the spiral interleaf angle by the golden-ratio angle $360^\circ \cdot 2/(\sqrt{5}+1) \approx 222.4969^\circ$ at every repetition time (TR) (see Fig. 1b). The unaliased FOV was defined as the reciprocal of the maximum sample spacing in k-space.

In Vivo Experiments

MRI experiments were performed on a commercial 1.5 T scanner (Signa Excite HD, GE Healthcare, Waukesha, WI). A body coil was used for radio frequency (RF) transmission, and a custom 4-channel upper airway receive coil array was used for RF signal reception. The receiver bandwidth was set to $\pm 125 \text{ kHz}$ (i.e., $4 \mu\text{s}$ sampling rate). One subject was scanned in supine position after providing informed consent in accordance with institutional policy.

A midsagittal scan plane of the upper airway was imaged using custom real-time imaging software (14). The spiral trajectory design followed those described in the Simulation section. The imaging protocol was: slice thickness = 5 mm, TR = 6.164 ms, temporal resolution = 80.1 ms. The golden-ratio view order scheme was compared with the

conventional bit-reversed 13-interleaf UDS scheme with all other imaging and scan parameters fixed (e.g., scan plane, shim and other calibrations, etc.). The volunteer was instructed to repeat “go pee shop okay bow know” for both the conventional bit-reversed UDS and golden-ratio acquisitions. The speech rate was maintained using a 160 bpm metronome sound that was communicated to the subject using the scanner intercom.

In conventional bit-reversed 13-interleaf UDS data, gridding reconstructions were performed using temporal windows of 8-TR and 13-TR. In golden-ratio spiral data, gridding reconstructions were performed using temporal windows of 8-TR, 13-TR, 21-TR, and 34-TR. Gridding reconstructions were based on interpolating the convolution of density compensated spiral k-space data with a 6×6 Kaiser-Bessel kernel onto a two-fold oversampled grids followed by taking 2D inverse fast Fourier transform (FFT) and deapodization (17). Root sum-of-squares (SOS) reconstruction from the 2 anterior elements of the coil was performed to obtain the final images. For comparison of images reconstructed from different temporal windows, image frames were reconstructed from the data in which the centers of each temporal window were aligned.

Blockwise Temporal Resolution Selection

In golden-ratio datasets, multiple temporal resolution videos can be produced retrospectively. We sought a procedure for automatic selection of the temporal window that was appropriate for each image region in each time frame, and the ability to use this to synthesize a single video.

We used time difference energy (*TDE*), as described in Eq. [1], as an indicator of motion. This was calculated for each block B_j and each time t :

$$TDE(B_j, t) = \sum_{t'=-T}^T \sum_{(x,y) \in B_j} |I(x, y, t) - I(x, y, t - t')|^2, \quad \text{Eq. [1]}$$

where $I(x, y, t)$ is image intensity at pixel location (x, y) and time t , and $2T$ is the number of adjacent time frames that are considered.

Temporal resolution selection was performed based on the alias-free high temporal resolution frames. We used sensitivity encoding (SENSE) reconstructed frames from 8-TR temporal resolution data (i.e., 49.3 ms temporal resolution) for the calculation of *TDE*. In addition, SENSE reconstructions were performed at each frame from 13-TR, 21-TR, and 34-TR temporal windows, whose corresponding temporal resolutions were 80.1 ms, 129.4 ms, and 209.6 ms. Data from all 4 elements of the coil were considered for the reconstruction. Alias-free coil sensitivity maps were obtained from 34-TR data. Sensitivity maps for each coil element were obtained by dividing the image at each element by the root SOS image of all elements. Frames were updated at every 4-TR = 24.7 ms (i.e., 40.6 frames per second). The spiral SENSE reconstruction was based on a non-linear iterative conjugate gradient algorithm with a total variation regularizer (18,19). Total variation regularization was effective at removing image noise while preserving high contrast signals such as the air-tissue boundaries. Iterations were terminated at the 20th iterate after visual inspection.

Intensity correction was performed at each frame using a thin plate spline fitting method (20). The SENSE reconstructed frames were first cropped to a 64×64 size that only contained the vocal tract regions of interest and then were interpolated to a 128×128 size in order to avoid the blockiness of the images. An 8×8 block and $T = 2$ was used to calculate *TDE*. Two spatially adjacent blocks were overlapped by 4 pixels in either the vertical or horizontal direction. The calculated *TDE* at each block was assigned to the central 4×4

block. TDE was normalized through the entire time frames. Temporal resolution selection at each 4×4 block was performed after a simple thresholding of the normalized TDE (TDE_{norm}). For $TDE_{norm} \geq 0.6$, 8-TR SENSE was assigned. For $0.4 \leq TDE_{norm} < 0.6$, 13-TR SENSE was assigned. For $0.2 \leq TDE_{norm} < 0.4$, 21-TR SENSE was assigned. For $TDE_{norm} < 0.2$, 34-TR SENSE was assigned. These settings were chosen empirically based on quality of the final synthesized video.

RESULTS

Figure 2 contains a plot of the unaliased FOV as a function of temporal resolution (i.e., the number of TRs) when retrospectively selecting a temporal window. For 13-TR, the 13-interleaf UDS supports a larger unaliased FOV than the golden-ratio view order. When the number of TRs becomes a Fibonacci number (e.g. 2,3,5,8,13,21,34), there is a change in the unaliased FOV for the golden-ratio method. Note the sudden increase in the unaliased FOV from 17.1 cm to 27.6 cm when the number of TRs changes from 20 to 21. For 8-TR, the 13-interleaf UDS provides inconsistent unaliased FOVs, which are 6.7 cm and 10 cm. The unaliased FOVs for the 13-interleaf UDS are smaller than those for the golden-ratio when the number of TRs is between 8 and 12. The golden-ratio view order provides a consistent unaliased FOV at every time point and at any temporal window chosen.

Figure 3 contains the images reconstructed from the data acquired when the subject was stationary. Note that the midsagittal slice of interest has a regional support of roughly 38 cm and has significant intensity shading due to coil sensitivity. The images reconstructed from the 13-interleaf UDS data have spatial aliasing artifacts in the regions posterior to the pharyngeal wall from coil 1 and in the regions superior to the hard palate and velum from coil 2. The root SOS image in Fig. 3(c) contains little or no aliasing artifacts within the vocal tract region of interest (denoted by the dashed box). Although the unaliased FOV (i.e., 20 cm) from the 13-interleaf UDS is larger than that from the golden-ratio sampling for the choice of 13-TR, aliasing artifacts indicated by white arrows in Fig. 3(c) are more prominent than in Fig. 3(d). The spatial aliasing pattern can be understood by examining the point spread functions (PSF) for each sampling pattern. The PSF from the 13-TR UDS had a ratio of maximum sidelobe to mainlobe peak (PSF_{max-sl}) of 0.061 and exhibited single sidelobe ring with a radius of 20 cm. The PSF from the 13-TR golden-ratio sampling had a PSF_{max-sl} of 0.044 and showed less coherent pattern with multiple sidelobe rings and lower sidelobe amplitude than the 13-TR UDS. Unlike the 13-interleaf UDS acquisition, a selection of long temporal window provides a large unaliased FOV in the golden-ratio acquisition. Note that aliasing artifacts are removed in the entire image from a selection of 34-TR window in Fig. 3(f).

Figure 4 contains image frames and time-varying intensity profiles from the conventional bit-reversed 13-interleaf UDS and golden-ratio methods when the subject produced the speech utterance “bow know”. Frames were updated at every TR. As seen in the undersampled 8-TR case of Fig. 4(c) and Fig. 4(d), the 13-interleaf UDS method produces aliasing artifacts that are periodic in time while the golden-ratio method produces less coherent aliasing in time. This periodicity in the aliasing is attributed to the inconsistent FOVs from the conventional bit-reversed 13-interleaf UDS as shown in Fig. 2. As seen from the 13-TR case in Fig. 4(c) and 4(d), the level of aliasing is higher for the golden-ratio result. Figure 4(d) shows that the intensity profile from the 8-TR result exhibits the sharpest transition of tongue tip motion (compare the black arrows in 8-TR, 13-TR, and 21-TR results).

The region-based temporal resolution selection method required the use of multiple temporal resolution videos reconstructed from iterative SENSE reconstructions, which substantially

increased computation time. The generation of 160 dynamic frames of SENSE reconstructions from 8-TR, 13-TR, 21-TR, and 34-TR took approximately 45 min, 49 min, 53 min, and 60 min, respectively, with a 3.06 GHz CPU and 3.48 GB RAM. The blockwise temporal resolution selection algorithm took approximately 2 min.

Figure 5 shows a result of the blockwise temporal resolution selection from the SENSE reconstructed golden-ratio framesets. The synthesized frames in Fig. 5(c) exhibit good assignment of four distinct temporal resolution videos. Less tongue tip blurring is seen as indicated by the yellow hollow arrows in Fig. 5(c) and 5(d) than the 34-TR result (see the red hollow arrow in Fig. 5(e)). A better visualization of the velum opening is seen as indicated by the yellow solid arrows in Fig. 5(c) and 5(e) than the 8-TR result (see the red solid arrow in Fig. 5(d)).

DISCUSSION

A new acquisition scheme that adopts a spiral golden-ratio view order has been demonstrated as a means to provide flexibility in retrospective selection of temporal resolution. The golden-ratio scheme has been compared with conventional bit-reversed 13-interleaf UDS acquisition, which is routinely used in our real-time speech MRI data collection at the University of Southern California. The spiral golden-ratio view order provides larger and consistent unaliased FOV when undersampling real-time data for higher temporal resolution. In addition, spiral interleaves are evenly distributed for any choices of the number of spiral interleaves at any time point, and hence a parallel imaging reduction factor can be flexibly chosen and applied to dynamic golden-ratio data. Auto-calibration with high resolution full FOV coil sensitivity maps is possible at any time point by utilizing fully-sampled temporal window data centered on that time point.

The proposed region-based temporal resolution selection method has limitations. The 8-TR (i.e., 49.3 ms temporal window) SENSE reconstructed frames served as a guide to select proper temporal resolution. However, they inherently lacked in temporal resolution and contained low image SNR. The velum and pharyngeal wall suffered from much lower SNR due to low coil sensitivity and potentially due to high parallel imaging g -factor. This can result in higher TDE regardless of motion. In addition, we performed SENSE reconstruction from the data whose temporal window is smaller than 8-TR, but resulting SENSE images produced inadequate image quality with significantly low SNR or blurred air-tissue boundaries with the use of a large regularization parameter. Higher acceleration with sufficient SNR may be possible by the use of a highly sensitive upper airway receive coil with higher channel counts (21).

The motivation for using blockwise processing was based on the following. First, air-tissue boundaries within a block typically move with similar speed. Second, blockwise processing helps to stabilize the calculation of TDE in the presence of noise. There is a trade-off in selection of the block size. For example, the choice of larger block size can result in improper assignment of temporal resolution for a block within which motion is not at uniform speed. The choice of smaller block size causes TDE to be more sensitive to noise.

The temporal resolution assignment procedure does not provide a strong link between the needed temporal resolution for an event and choice of retrospective temporal resolution. An additional navigator sequence may help to obtain the required temporal resolution information although it reduces scan efficiency. Ref. (9) by Tasko and McClean reports that the speed of the tongue tip during fluent speech was measured using electromagnetic articulometer (EMA) and was up to 200 mm/s. With the 3 mm spatial and 49 ms temporal resolution from 8-TR SENSE, it is anticipated that more than 3 pixels will experience

temporal blurring around the tongue tip region of interest if the speed is 200 mm/s. Hence, higher spatio-temporal resolution frames will be necessary for estimating temporal bandwidth more reliably.

Recent real-time spiral speech MRI has been demonstrated with 80~100 ms temporal resolution (3,5,13), but it lacks in temporal resolution compared to other speech imaging technologies such as EMA and ultrasound. Higher temporal resolution can be achieved by lowering spatial resolution or designing longer spiral readout with a fewer number of interleaves. Lower spatial resolution imaging may lose details of fine structures such as the epiglottis or lead to more difficulties in resolving the narrowing in the airway, e.g., the constriction between the alveolar ridge and tongue tip in certain sound productions such as the fricative /s/. Lengthening the spiral readout would cause images to be more susceptible to blurring or distortion in the air-tissue boundaries due to a large amount of resonance offset from air-tissue magnetic susceptibility. Correction of blurring or distortion artifacts is challenging in real-time upper airway MRI because of difficulty in estimating accurate field map. Effective off-resonance correction from real-time golden-ratio view order data is an interesting area for investigation. Alternatively, real-time radial speech MRI with 55 ms temporal resolution has been demonstrated using a short-TR radial fast gradient echo sequence and parallel image reconstruction in combination with temporal filtering (22). Improved temporal resolution imaging may be achieved by adopting an additional navigator sequence in the acquisition and a spatiotemporal model in the reconstruction (2,23).

We have focused on an example of nasal sound production study in which knowledge of the timing of oral and velar coordination is important for modeling temporal changes in the constriction degrees of articulators under different syllable contexts (13). This particular articulation involves a rapid tongue tip motion and relatively slow velum motion when producing nasal consonant /n/, and is well suited for investigating the importance of flexibility in temporal resolution selection. Similarly, the golden-ratio method can be applied to other articulatory timing studies which are investigated in the literature (24).

The acoustic noise generated by the MRI gradients during conventional bit-reversed 13-interleaf UDS imaging is temporally periodic. This can be exploited for high quality adaptive noise cancellation and audio recordings during speech production in the magnet (6). One difficulty with golden-ratio imaging is that the MRI gradient noise is no longer periodic. Advanced models are therefore required for acoustic noise cancellation during golden-ratio spiral imaging, and remain as future work.

CONCLUSION

We have demonstrated the application of a spiral golden-ratio temporal view order to imaging a midsagittal slice of the vocal tract during fluent speech. Simulation studies showed that the golden-ratio method provided larger and consistent unaliased FOV when retrospectively undersampling real-time data than the conventional bit-reversed 13-interleaf uniform density spiral. In nasal speech imaging studies, the proposed method provided an improved depiction of rapid tongue tip movement with less temporal blurring and velum lowering with higher SNR and potentially reduced aliasing artifacts. The region-based temporal resolution selection method synthesizes a single video from multiple temporal resolution videos available in the golden-ratio real-time data and potentially facilitates subsequent vocal tract shape analysis.

Acknowledgments

NIH R01-DC007124.

The authors acknowledge the support and collaboration of the Speech Production and Articulation kNOWLEDge (SPAN, <http://sail.usc.edu/span>) group at the University of Southern California. The authors would like to thank Stefanie Remmele for valuable discussions on the golden-ratio sampling.

References

1. Demolin D, Hassid S, Metens T, Soquet A. Real-time MRI and articulatory coordination in speech. *Comptes Rendus Biologies*. 2002; 325(4):547–556. [PubMed: 12161933]
2. Sutton BP, Conway C, Bae Y, Brinegar C, Liang Z-P, Kuehn DP. Dynamic imaging of speech and swallowing with MRI. *IEEE EMBS*. 2009:6651–6654.
3. Narayanan S, Nayak K, Lee S, Sethy A, Byrd D. An approach to real-time magnetic resonance imaging for speech production. *J Acoust Soc Am*. 2004; 115(4):1771–1776. [PubMed: 15101655]
4. Bresch E, Kim YC, Nayak KS, Byrd D, Narayanan SS. Seeing speech: Capturing vocal tract shaping using real-time magnetic resonance imaging. *IEEE Signal Processing Magazine*. 2008:123–132.
5. Bresch E, Narayanan SS. Region segmentation in the frequency domain applied to upper airway real-time magnetic resonance images. *IEEE Trans Medical Imaging*. 2009; 28:323–338.
6. Bresch E, Nielsen J, Nayak KS, Narayanan S. Synchronized and noise-robust audio recordings during real-time MRI scans. *J Acoust Soc Am*. 2006; 120(4):1791–1794. [PubMed: 17069275]
7. Proctor M, Goldstein L, Byrd D, Bresch E, Narayanan S. Articulatory comparison of Tamil liquids and stops using real-time magnetic resonance imaging. *J Acoust Soc Am*. 2009; 125(4):2568.
8. Adams SG, Weismer G, Kent RD. Speaking rate and speech movement velocity profiles. *J Speech and Hear Res*. 1993; 36:41–54. [PubMed: 8450664]
9. Tasko SM, McClean MD. Variations in articulatory movement with changes in speech task. *J Speech Lang Hear Res*. 2004; 47(1):85–100. [PubMed: 15072530]
10. Ostry DJ, Munhall KG. Control of rate and duration of speech movements. *J Acoust Soc Am*. 1985; 77(2):640–648. [PubMed: 3882804]
11. Browman, CP.; Goldstein, L. Tiers in articulatory phonology, with some implications for casual speech. In: Kingston, J.; Beckman, ME., editors. *Papers in laboratory phonology I: Between the grammar and physics of speech*. 1990. p. 341-376.
12. Winkelmann S, Schaeffter T, Koehler T, Eggers H, Doessel O. An optimal radial profile order based on the golden ratio for time-resolved MRI. *IEEE Trans Medical Imaging*. 2007; 26(1):68–76.
13. Byrd D, Tobin S, Bresch E, Narayanan SS. Timing effects of syllable structure and stress on nasals: a real-time MRI examination. *Journal of Phonetics*. 2009; 37:97–110. [PubMed: 20046892]
14. Santos JM, Wright GA, Pauly JM. Flexible real-time magnetic resonance imaging framework. *Proceedings of the 26th Annual Meeting of IEEE EMBS*. 2004; 47:1048–1051.
15. Spielman DM, Pauly JM, Meyer CH. Magnetic resonance fluoroscopy using spirals with variable sampling densities. *Magn Reson in Med*. 1995; 34(3):388–394. [PubMed: 7500878]
16. Nayak KS, Pauly JM, Kerr AB, Hu BS, Nishimura DG. Real-time color flow MRI. *Magn Reson Med*. 2000; 43:251–258. [PubMed: 10680689]
17. Jackson J, Meyer CH, Nishimura D, Macovski A. Selection of convolution function for Fourier inversion using gridding. *IEEE Trans Med Imaging*. 1991; 10:473–478. [PubMed: 18222850]
18. King, KF. Combined compressed sensing and parallel imaging. *Proceedings of the 16th Annual Meeting of ISMRM; Toronto*. 2008. (abstract 1488)
19. Lustig M, Donoho D, Pauly JM. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn Reson Med*. 2007; 58:1182–1195. [PubMed: 17969013]
20. Liu C, Bammer R, Moseley ME. Parallel imaging reconstruction for arbitrary trajectories using k-space sparse matrices (kSPA). *Magn Reson in Med*. 2007; 58:1171–1181. [PubMed: 17969012]
21. Hayes, CE.; Carpenter, C.; Evangelou, IE.; Chi-Fishman, G. Design of a highly sensitive 12-channel receive coil for tongue MRI. *Proceedings of the 15th Annual Meeting of ISMRM; Berlin*. 2007. (abstract 449)
22. Uecker M, Zhang S, Voit D, Karaus A, Merboldt K-D, Frahm J. Real-time MRI at a resolution of 20 ms. *NMR in Biomed*. 2010; 23(8):986–994.

23. Liang, Z-P. Spatiotemporal imaging with partially separable functions. International Symposium on Biomedical Imaging; 2007. p. 988-991.
24. Marin S, Pouplier M. Temporal organization of complex onsets and codas in American English: Testing the predictions of a gestural coupling model. Journal of Motor Control. 2010; 14(3):380–407.

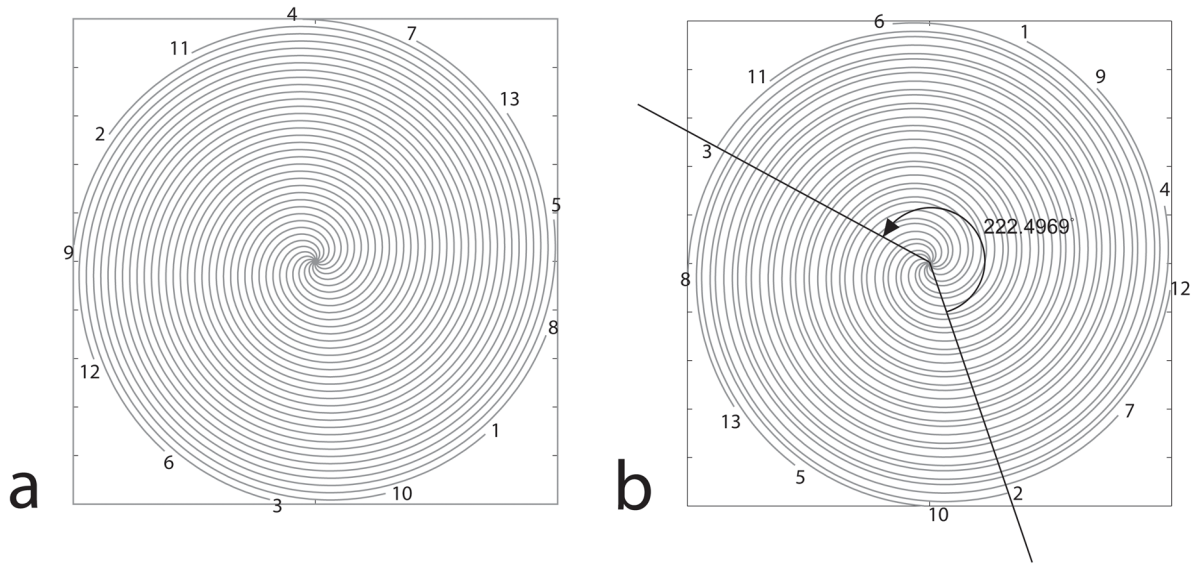


Figure 1.

k-space trajectories for (a) conventional bit-reversed 13-interleaf UDS and (b) golden-ratio spiral view order when samples from 13 consecutive TRs are combined. Temporal view orders are marked with the numbers on each end of the spiral interleaves. In (a), the angle spacing between spatially adjacent spiral interleaves is uniform with an angle of $360^\circ/13$. In (b), the angle spacing between spatially adjacent spiral interleaves is not uniform but the angle increment in successive view numbers is uniform with an angle of $360^\circ \cdot 2/(\sqrt{5}+1) \approx 222.4969^\circ$.

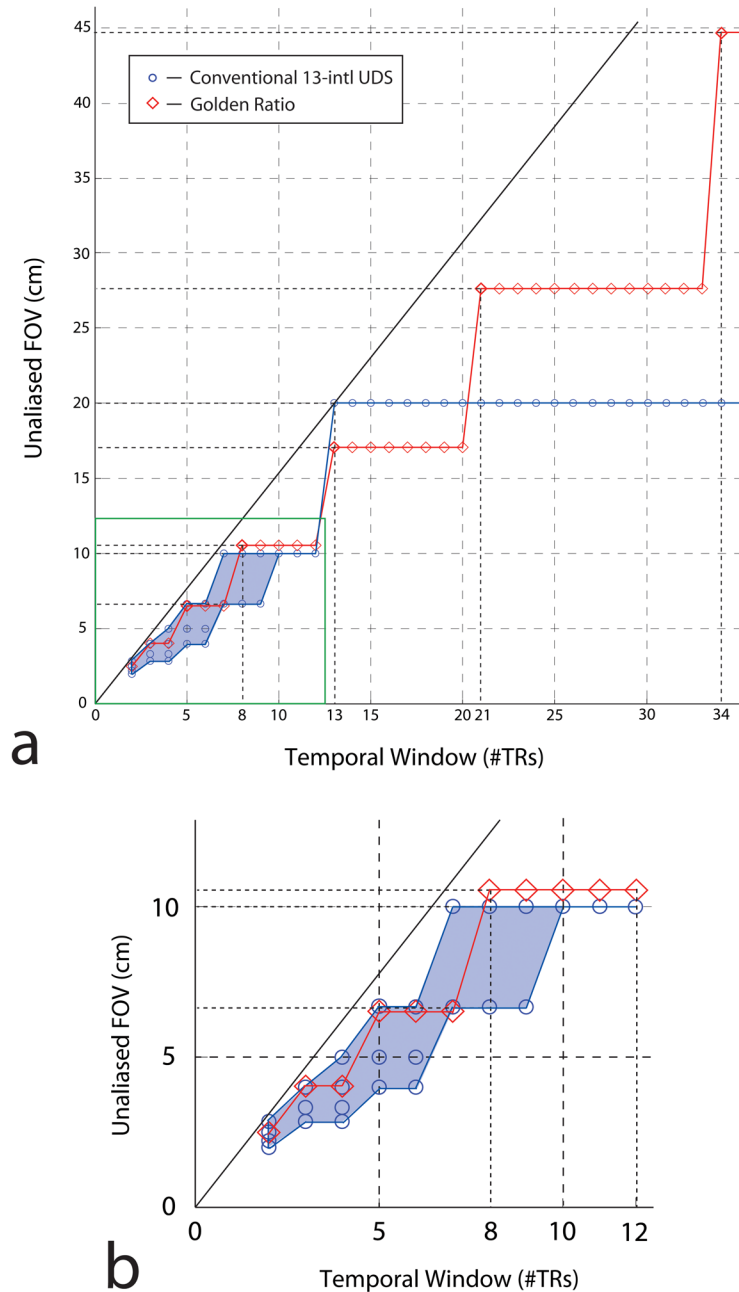


Figure 2. Retrospective selection of temporal resolution: **(a)** Comparison of unaliased FOV between the golden-ratio view order and conventional bit-reversed 13-interleaf UDS sampling. **(b)** The enlargement of the region within the green rectangle in (a). The blue shaded region in (a,b) indicates that unaliased FOV varies in conventional bit-reversed 13-interleaf UDS when the number of TRs is less than 10. The black solid line illustrates a linear relationship between unaliased FOV and temporal resolution when UDS trajectories are designed under the constraints of the same spatial resolution and readout duration. Note that the unaliased FOV is fixed as 20 cm for a temporal window length (13-TR) for the conventional bit-reversed 13-interleaf UDS, but it increases with temporal window length for the golden-ratio view order.

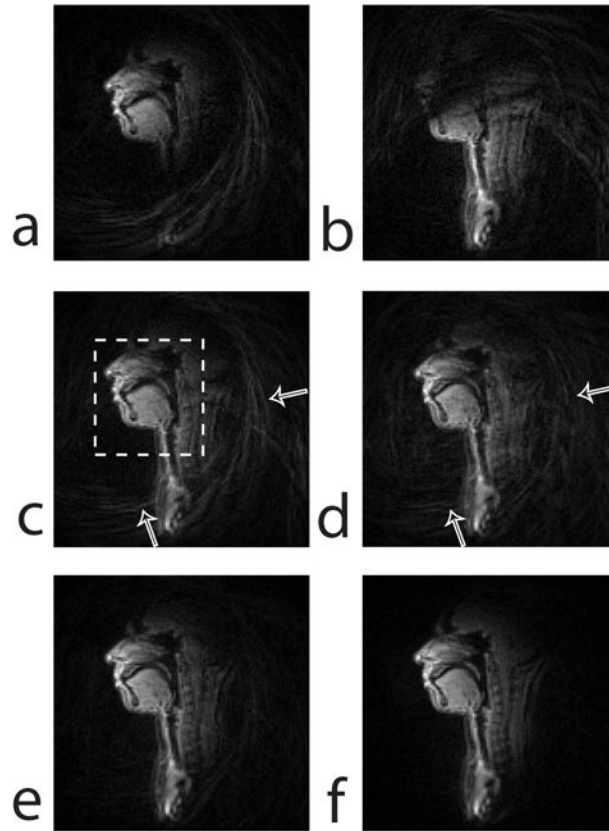


Figure 3. Midsagittal images with a large reconstruction field-of-view (FOV) of $38 \times 38 \text{ cm}^2$ reconstructed from the data acquired in static posture. **(a–c)** Conventional bit-reversed 13-interleaf UDS. (a) image from coil 1, (b) image from coil 2, (c) root sum-of-squares (SOS) of the coil 1 and coil 2 images. The region within the dashed box in (c) is the vocal tract regions of interest (ROIs). For speed-up of the spiral acquisition, the FOV of the 13-interleaf UDS is typically chosen to be small such that aliasing artifacts are not observed in the vocal tract ROIs. **(d–f)** Root SOS of the coil 1 and coil 2 images reconstructed from data acquired using the spiral golden-ratio acquisition: reconstruction from (d) 13-TR, (e) 21-TR, and (f) 34-TR data. Spatial aliasing artifacts are completely removed in (f) because of larger FOV available from the golden-ratio method. The image SNR in (f) is higher than that in (d) and (e).

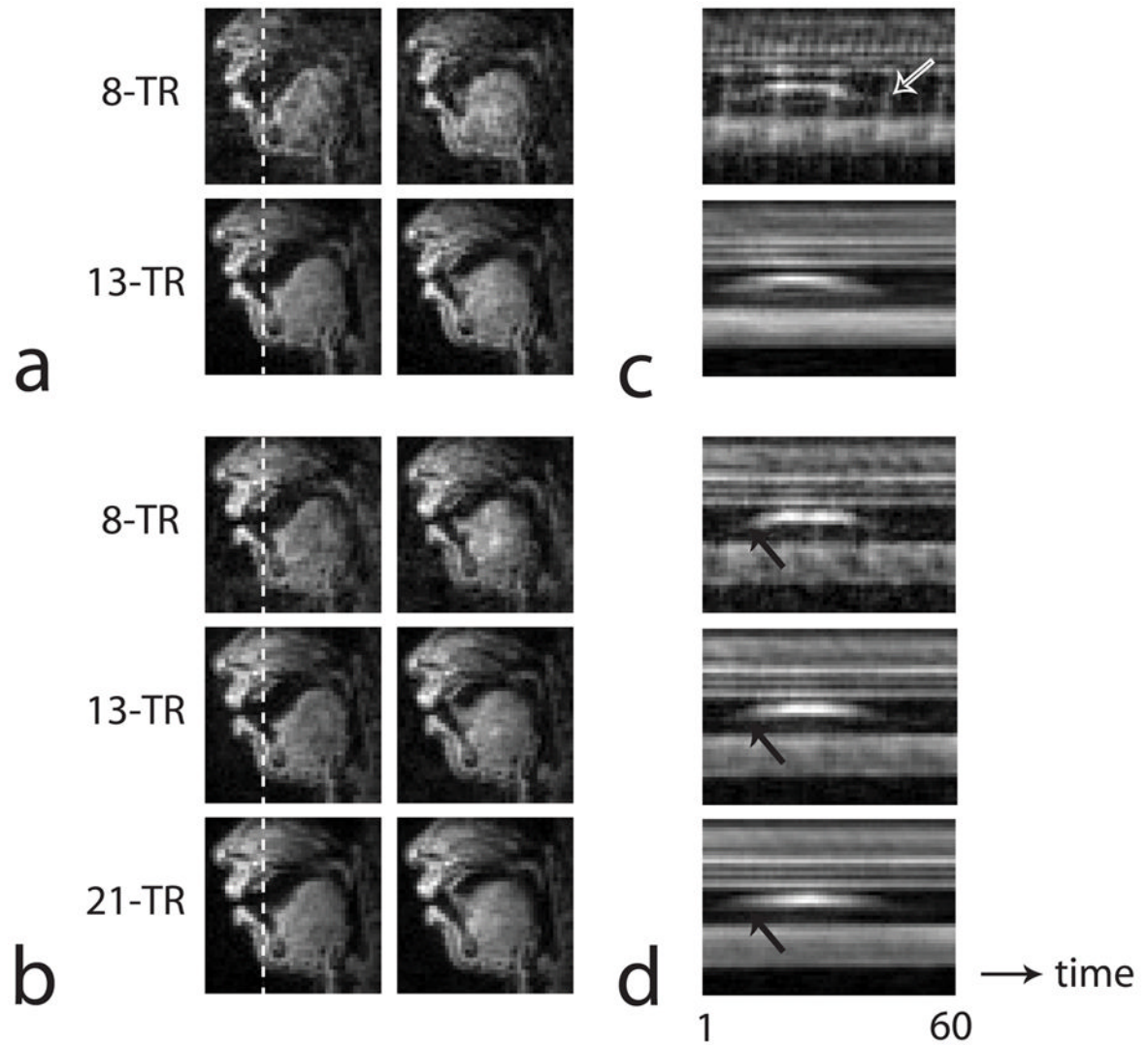


Figure 4.

Gridding reconstructed dynamic frames and time intensity profiles from (a,c) bit-reversed 13-interleaf uniform density spiral data and (b,d) spiral golden-ratio view order data. A 40×40 matrix containing only the vocal tract region of interest is shown. Retrospective selection of temporal resolution is performed using 8-TR and 13-TR in (a), and 8-TR, 13-TR, and 21-TR in (b). Frame update rate was $1\text{-TR} = 6.164$ ms. Two example frames (frame 1, 6) are shown (frame 1 is relatively stationary and frame 6 is captured during rapid tongue tip motion). Time intensity profiles from the image column indicated by the dashed lines are shown for (c) bit-reversed 13-interleaf UDS and (d) golden-ratio view order data.

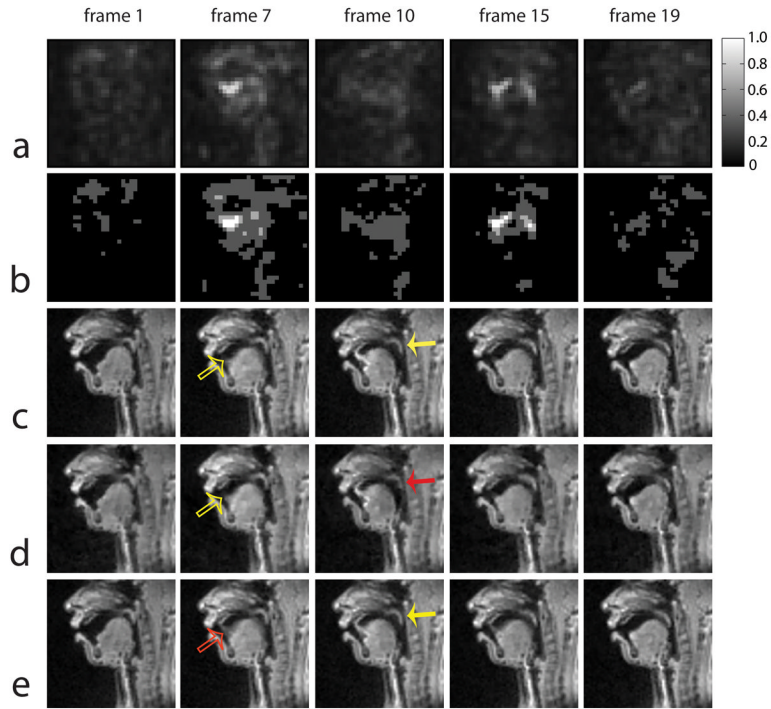


Figure 5. Blockwise temporal resolution selection and synthesis of a single video from four temporal resolution videos. Five representative frames are shown that are captured when the subject produced /bono/. **(a)** Normalized time difference energy (*TDE*) map. **(b)** Temporal resolution selection map [white: 49 ms (= 8-TR) temporal resolution, bright gray: 80 ms (= 13-TR) temporal resolution, dark gray: 129 ms (= 21-TR) temporal resolution, black: 210 ms (= 34-TR) temporal resolution]. **(c)** Synthesized temporal resolution frames based on the temporal resolution selection map in (b). **(d)** 49 ms (= 8-TR) temporal resolution frames. **(e)** 210 ms (= 34-TR) temporal resolution frames.