IMAGING METHODOLOGY – Notes

Magnetic Resonance in Medicine 78:2275–2282 (2017)

Feasibility of Through-Time Spiral Generalized Autocalibrating Partial Parallel Acquisition for Low Latency Accelerated Real-Time MRI of Speech

Sajan Goud Lingala,[1]* Yinghua Zhu,[1] Yongwan Lim,[1] Asterios Toutios,[1] Yunhua Ji,[2] Wei-Ching Lo,[3] Nicole Seiberlich,[3] Shrikanth Narayanan,[1] and Krishna S. Nayak[1]

Purpose: To evaluate the feasibility of through-time spiral generalized autocalibrating partial parallel acquisition (GRAPPA) for low-latency accelerated real-time MRI of speech.

Methods: Through-time spiral GRAPPA (spiral GRAPPA), a fast linear reconstruction method, is applied to spiral (k-t) data acquired from an eight-channel custom upper-airway coil. Fully sampled data were retrospectively down-sampled to evaluate spiral GRAPPA at undersampling factors $R = 2$ to 6. Pseudo-golden-angle spiral acquisitions were used for prospective studies. Three subjects were imaged while performing a range of speech tasks that involved rapid articulator movements, including fluent speech and beat-boxing. Spiral GRAPPA was compared with view sharing, and a parallel imaging and compressed sensing (PI-CS) method.

Results: Spiral GRAPPA captured spatiotemporal dynamics of vocal tract articulators at undersampling factors ≤ 4. Spiral GRAPPA at 18 ms/frame and 2.4 mm^2/pixel outperformed view sharing in depicting rapidly moving articulators. Spiral GRAPPA and PI-CS provided equivalent temporal fidelity. Reconstruction latency per frame was 14 ms for view sharing and 116 ms for spiral GRAPPA, using a single processor. Spiral GRAPPA kept up with the MRI data rate of 18ms/frame with eight processors. PI-CS required 17 minutes to reconstruct 5 seconds of dynamic data.

Conclusion: Spiral GRAPPA enabled 4-fold accelerated real-time MRI of speech with a low reconstruction latency. This approach is applicable to wide range of speech RT-MRI experiments that benefit from real-time feedback while visualizing rapid articulator movement. **Magn Reson Med 78:2275–2282, 2017. © 2017 International Society for Magnetic Resonance in Medicine.**

Key words: real-time MRI; low latency imaging; speech production

INTRODUCTION

Real-time MRI (RT-MRI) is a powerful tool to safely assess and quantify the vocal tract dynamics during speech production (1–3). It has several compelling advantages over alternate modalities (X-rays, ultrasound, and electromagnetic articulography), including the lack of any ionizing radiation, ability to image arbitrary scan planes, and visualization of deep soft-tissue structures such as the velum, epiglottis, larynx, and pharyngeal wall. MRI has become a useful tool in human speech production research and has allowed scientists to address open questions in the areas of phonetics, phonology, language acquisition, and language disorders (1–3). RT-MRI of upper-airway dynamics has also shown value in clinical assessment of velopharyngeal insufficiency (4,5), cleft palate (6,7), treatment planning and post-treatment functional evaluation of speech and swallowing (8), and dynamic assessment of airway collapses in obstructive sleep apnea (9).

RT-MRI is challenged by trade-offs between the achievable spatial resolution, temporal resolution, slice coverage, and signal to noise. Several rapid MRI methods based on non-Cartesian imaging, parallel imaging, and compressed sensing (CS) have been applied to improve the above trade-offs (10–16). The latency time, which is the time between acquisition of a set of raw data and reconstruction of the final image, is a useful criterion to classify the above methods to either on-the-fly or off-line methods. On-the-fly methods are those which achieve minimal latency requirements for a given application, whereas off-line methods are those that have larger than tolerable latencies. The minimum latency time requirements are typically dictated by the application at hand and can be in the range of 100 to 1,000 ms. For instance, latencies of 500 to 1,000 ms can be tolerated during interactive localization (17,18), whereas latencies of up to 330 ms have been reported in interventional guidance (19). It is also worth mentioning that in non-MRI applications, latencies of up to 200 ms have been established for audio in conversational speech (human mouth to ear) (20) and up to 100 to 1,000 ms in interactive gaming applications (21).

On-the-fly methods that combine non-Cartesian imaging with view sharing have been used in several studies (1,10,12). Bresch et al and Narayanan et al (1,10) used short spiral readouts at 1.5 Tesla (T) to visualize speech at native time resolutions of 54 to 78 ms and spatial resolution of 3.0 to 2.4 mm^2. Interactive imaging is desired in scenarios that require real-time feedback, such as

[1]Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, California, USA.

[2]Department of Biomedical Engineering, University of Southern California, Los Angeles, California, USA.

[3]Biomedical Engineering, Case Western Reserve University, Cleveland, Ohio, USA.

*Correspondence to: Sajan Goud Lingala, PhD, 3740 McClintock Avenue, Ming Hsieh Department of Electrical Engineering (EEB 413), University of Southern California, Los Angeles, CA 90089. E-mail: lingala@usc.edu

Received 1 August 2016; revised 8 December 2016; accepted 27 December 2016

DOI 10.1002/mrm.26611

Published online 10 February 2017 in Wiley Online Library (wileyonlinelibrary.com).

immediate visualization of articulatory movements in various planes to localize a desired slice, immediate visualization, and correction of artifacts (eg, reduction of off-resonance blurring in spiral imaging by on-the-fly adjustment of center frequency).

Iterative CS methods that exploit spatiotemporal redundancies of dynamic images have demonstrated greater acceleration levels (>6- to 7-fold). These invoke constraints based on low-rank assumptions, transform sparsity, or both (11–16,22). These have demonstrated improved spatiotemporal resolutions in 2D RT-MRI (up to 1.5–2.4 mm^2 and 10–33 ms/frame) and also more recently in 3D RT-MRI (frame rate of 166 frames/sec) (22). The reconstructions in these methods involve nonlinear optimization, rely on the availability of an adequate number of temporal frames for efficient exploitation of redundancies, and are largely implemented off-line. These methods are challenged by complexity and computation times. However, several approaches that leverage parallelism and graphical processing units have been able to achieve latencies up to 47.5 ms/frame in Sorensen et al (23), 87 ms/frame in Schaetz et al (24), 90 ms/frame in Uecker et al (25), and 300 ms/frame in Smith et al (26) for similar applications.

In this work, the feasibility of a fast linear k-space-based reconstruction method through-time spiral generalized autocalibrating partial parallel acquisition (GRAPPA; spiral GRAPPA) (27,28) is evaluated for accelerated RT-MRI of speech with low latency and robust image quality. Through-time non-Cartesian GRAPPA uses multiple fully sampled time frames in a one-time calibration phase to learn the k-space geometry-specific GRAPPA weights of a given kernel. With the known weights, it performs a frame-by-frame linear reconstruction to estimate the missing k-space samples in every coil. Through-time non-Cartesian GRAPPA has been applied and evaluated in several dynamic and relaxometry MRI applications, including real-time cardiac MRI (29), renal MRI angiography (30), liver perfusion MRI (31), and T_1 mapping (32). Here, we evaluate it to efficiently exploit the acceleration capabilities offered by a custom eight-channel upper-airway coil and spiral trajectories for dynamic upper-airway imaging. Through quantitative and qualitative evaluation in retrospective undersampling experiments of fully sampled data, we determine the length of calibration data needed for GRAPPA weight estimation and also evaluate the spatiotemporal fidelity of spiral GRAPPA reconstructions at various undersampling factors (R). A pseudo-golden-angle spiral sequence with linear view ordering in the calibration phase is used, with a golden angle ordering in the acquisition phase, as described in Kim et al (33). Comparisons are performed against existing on-the-fly view-sharing reconstruction and an off-line sparse SENSE based parallel imaging and compressed sensing method (PI-CS). Examples from 3 subjects are shown to visualize rapidly moving articulators with a range of speech tasks.

METHODS

Experiments were performed on a GE Signa Excite 1.5T scanner with a custom eight-channel upper-airway coil, which had four channels on either side of the jaw. The coil elements were arranged to offer diverse sensitivity in the superior-inferior and anterior-posterior directions. It provides high sensitivity over all upper-airway articulators, including the lips, tongue, velum, epiglottis, glottis, and pharyngeal wall (14).

Gradient echo–based multishot interleaved spiral sequences (field of view = 20 cm^2; slice thickness [Δz] = 6 mm; readout time [T_{read}] = 2.5 ms; repetition time [TR] = 6.004 ms; echo time = 0.8 ms; flip angle = 15^0; receive bandwidth = ±125 kHz) were implemented at different spatial resolutions, while making maximum use of gradients (40 mT/m amplitude and 150 mT/m/ms slew rate). All data were acquired by a real-time interactive imaging platform (RT-Hawk; Heart Vista Inc., Los Altos, CA, USA) (34).

In Vivo Validation Using Retrospective Down-Sampling

Data from 1 subject were acquired using the above multishot spiral sequence with linear interleaf order (spatial resolution = 2.5 mm^2; image matrix size: 80×80; 12 interleaves per fully sampled frame; angle increments of 30^0) and a midsagittal scan plane.

Two speech tasks were considered: 1) repetitions of counting numbers "one, two, three, four, five" at a normal pace, used for spiral GRAPPA calibration, and 2) repetitions of the phrase "ala-ara-asa-asha" at a normal pace, used as reference data. The reference data were retrospectively down-sampled at undersampling factors (R) = 2, 3, 4, and 6 by selecting every Rth interleave in each fully sampled frame. The missing interleaves for each coil were recovered by spiral GRAPPA. This retrospective selection of every Rth interleave in each fully sampled frame enabled evaluation of artifact behavior only from undersampling. Any possible motion blurring artifacts attributed to the relatively large temporal footprint were applied in a similar manner in both the fully sampled and retrospective undersampled data. Images were reconstructed from the non-Cartesian k-space data using the inverse nonuniform fast Fourier transform (nuFFT) (35) and the individual coils combined. Images with high signal to noise obtained from time-averaged data (with a temporal window of 5 seconds) was used for coil sensitivity map estimation with the eigen decomposition method (36). Spiral GRAPPA reconstructions were assessed as a function of number of time frames in the calibration data (t_{calib}) with a fixed block size of 1×1 in the readout and azimuthal directions. Reconstructions ($\hat{\gamma}(\mathbf{x}, t)$) were evaluated against the reference data ($\gamma_{ref}(\mathbf{x}, t)$) qualitatively in terms of spatiotemporal fidelity of moving articulator boundaries. Quantitative evaluation in the image domain using the average structural similarity measure index (SSIM) measure, as described in Wang et al (37), and the average normalized root mean square error metric (nRMSE) were also performed (Eq. 1):

$$\text{nRMSE} = \frac{1}{N}\sum_{i=1}^{N}\sqrt{\frac{||\gamma_{ref}(\mathbf{x}, t_i) - \hat{\gamma}(\mathbf{x}, t_i)||_2^2}{||\gamma_{ref}(\mathbf{x}, t_i)||_2^2}}; \qquad [1]$$

where N is the number of time frames. Note that any possible motion-blurring artifacts attributed to the 72 ms/frame temporal footprint in the reference data will be present in both the reference and retrospective

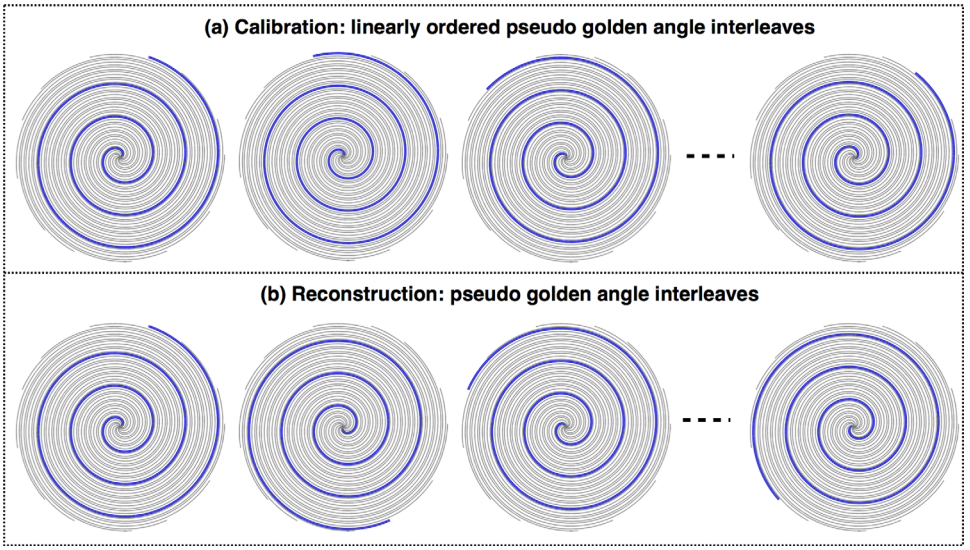

FIG. 1. k-t spiral sampling in prospective studies: pseudo-golden-angle sampling was considered with angle increments of 222.49^0 and a period of 13 interleaves. In the calibration stage (**a**), the interleaves are sorted in ascending order of the distribution. In the reconstruction stage (**b**), the interleaves are acquired with golden-angle increments.

undersampled data. The nRMSE therefore only quantifies the error attributed to undersampling.

Evaluation Using Pseudo-Golden-Angle View Order k-t Sampling

Data from two subjects were acquired using a pseudo-golden-angle multishot spiral sequence (spatial resolution = 2.4 mm^2; image matrix size: 84×84; 13 interleaves per fully sampled frame; angle increments of $\sim 222.49^0$; periodicity = 13 interleaves) (33). An angle increment of $\sim 222.49^0$ (or equivalently $360^0 - 222.49^0 = 139.51^0$) is used because we use conventional spiral trajectories as opposed to spiral in-out (or full-spoke radial) trajectories, where $\sim 111.246^0$ is applicable (33,38,39). Pseudo-golden-angle sampling refers to the scheme where the (k,t) sampling pattern repeats after a fixed number of golden-angle increments (33). In this work, the sampling scheme is repeated with periodicity of 13 interleaves because this corresponds to the number of interleaves for full sampling, and the calibration requires repetitions of several fully sampled frames. During calibration, data were acquired with a linear interleaf ordering from the pseudo-golden-angle distribution (see Fig. 1). A task of producing fluent speech (counting numbers at the subject's natural pace) was performed for 12 seconds during calibration data collection. The rationale for linearly sorting the interleaves is to ensure minimal object motion between successive k-space interleaves, which, in turn, allows for the GRAPPA weights to efficiently capture coil geometry dependencies while being robust to object motion (40).

Multiple (k-t) data sets were acquired as a part of a 90-minute scan session with a range of stimuli including production of beat-boxing sounds (41) and fluent speech (counting numbers at a rapid pace). The audio was simultaneously recorded using a fiberoptic microphone (Optoacoustics Ltd., Moshav Mazor, Israel) and custom recording and synchronization setup (42,43). Spiral GRAPPA reconstructions were performed after varying the number of interleaves per reconstructed frame, which allowed the same task to be visualized at different temporal resolutions. Specifically, 13 and 3 interleaves/frame were considered, which respectively corresponded to native time resolutions of 78 ms/frame, and 18 ms/frame.

Comparisons With Existing On-the-Fly and Off-Line Methods

Spiral GRAPPA is compared against existing on-the-fly gridding reconstruction with and without view sharing (1,10), and an off-line sparse sensitivity encoding–based PI-CS method that exploits temporal finite difference sparsity (14). View sharing is implemented by forming images from 13 interleaves/frame, with a step size of three interleaves. This corresponded to a native time resolution of 78 ms and a frame rate of 55 frames per second (fps). Spiral GRAPPA and PI-CS were implemented with 3 interleaves/frame, that is, with a native time resolution of 18 ms/frame, and frame rate of 55 fps. For view sharing and spiral GRAPPA, reconstruction was implemented in a frame-by-frame manner. For PI-CS, dynamic data of 5 seconds (ie, $\sim$277 image time frames) were reconstructed simultaneously using nonlinear conjugate gradient optimization. The calibration size in spiral GRAPPA was $1 \times 1 \times 150$ respectively in the read out, azimuthal, and temporal directions. To ensure overcompleteness, we have used more number of measurements than the required of required autocalibration signal points (an oversampling factor of $\sim$3.1). The total time for calibration was approximately 12 seconds. The regularization parameter in PI-CS was empirically adjusted to 0.1 based on a L-curve heuristic as previously described (14). Raw data were exported from the RT-HAWK platform, and all the reconstructions were implemented off-line within MATLAB (The MathWorks, Inc., Natick, MA, USA) on an Intel(R) Xeon(R) CPU E5-2698 v3; 2.30GHz; 40 MB of L3 cache; 16 processors. Qualitative assessment of spatiotemporal fidelity of the reconstructions was performed.

RESULTS

Figure 2 shows retrospective down-sampling spiral GRAPPA reconstructions at R = 4, where nRMSE and

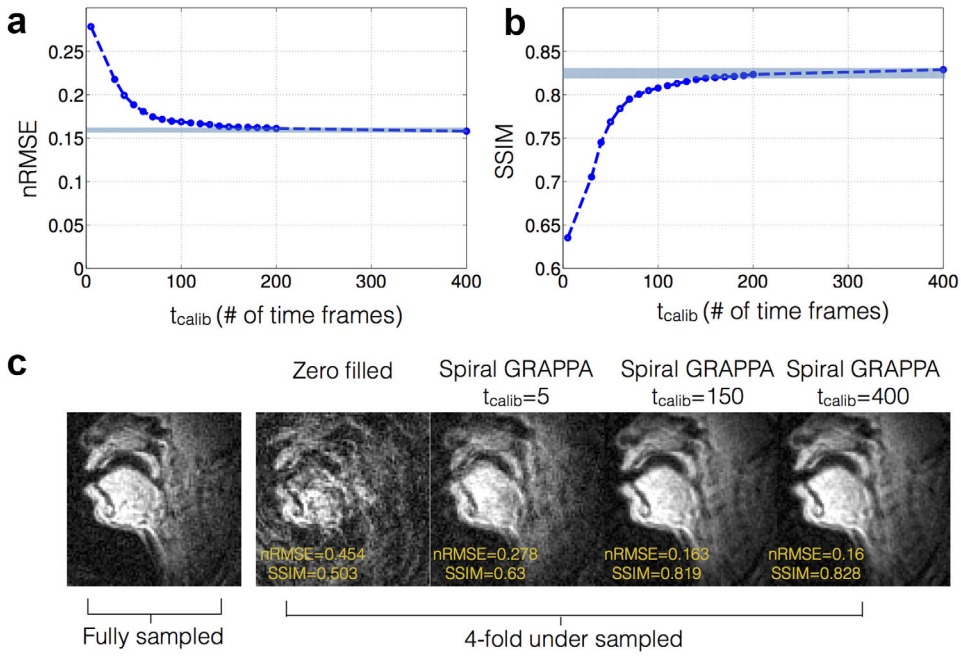

FIG. 2. Effect of calibration size on spiral GRAPPA reconstruction: (**a**) nRMSE; (**b**) SSIM between the reconstructions and fully sampled images is plotted v/s the number of time frames during calibration (t_{calib}). (**c**) Example spatial frame from reconstructions using fully sampled data, zero filled at $R=4$, spiral GRAPPA at $R=4$ with different t_{calib}. nRMSE and SSIM respectively decreased and increased monotonically with increasing t_{calib}, however plateaus at $t_{calib}=150$. Similar reconstructions were observed with $t_{calib} > 150$ (see blue highlighted area that corresponds to nMRSE within 0.3% and SSIM within 0.9%). In this work, we use $t_{calib}=150$, which is ~11.7 seconds.

SSIM are plotted as a function of t_{calib}. For low $t_{calib} < 10$, substantial noise amplification is seen. nRMSE and SSIM respectively decrease and increase monotonically with increasing t_{calib}, but saturates at $t_{calib} \geq 150$. It can be seen that there is no considerable difference in the reconstructions from $t_{calib}=150$ (nRMSE = 0.163; SSIM = 0.819) and $t_{calib}=400$ (nRMSE = 0.16; SSIM = 0.828). Based on these observations, we used $t_{calib}=150$ in the remaining experiments, which required approximately 11.7 seconds.

Figure 3 shows undersampled spiral GRAPPA reconstructions evaluated against a fully sampled reference reconstruction. The image time profiles and the error time profiles qualitatively highlight the spatiotemporal fidelity at various Rs. As expected, noise amplification is observed with increasing R (attributed to fewer interleaves), which was consistent with the factor $\sqrt{\frac{N_{fs}}{N_{us}}}$; where N_{fs} and N_{us} are respectively the number of acquired interleaves in the reference fully sampled data and undersampled data. Spiral GRAPPA is observed to maintain the spatiotemporal dynamics of the articulators well up to $R=3$ or 4 (nRMSE, <0.163; SSIM, >0.81), whereas at $R \geq 6$, spatial blurring artifacts appear at edges. Supporting Video S1 shows the corresponding dynamic movies.

Figure 4 compares results from a subject repating the phrase: "One-two-three-four-five" at a rapid pace. Zero filled nUFFT (Fig. 4a) provides substantial undersampling artifacts. View sharing (Fig. 4b) blurred the fast articulatory movements, such as tongue tip hitting the hard palate, attributed to the large native time resolution window of 78 ms. In comparison, spiral GRAPPA (Fig. 4c) provides a clear depiction of these fast-moving articulators, which is attributed to its superior time resolution of 18 ms.

Figure 5 compares the results from a subject producing beat-boxing sounds. These sounds involve rapid articulatory dynamics of the opening and closing of the lips, tongue tip motion, and velopharyngeal closures. View sharing (Fig. 5a) blurred the articulatory movements during rapid events, such as the inward movement of the tongue and velopharyngeal closures, both of which could be seen clearly in the PI-CS and spiral GRAPPA reconstructions. Compared to view sharing and spiral GRAPPA, PI-CS demonstrated higher signal-to-noise ratio attributed to inherent sparsity-based denoising and fewer undersampling artifacts; however, it may also introduce new nontrivial artifacts (3,44). Supporting Video S2 shows the dynamic movie of spiral GRAPPA along with noise-cancelled audio during free-style beat-boxing sounds.

View sharing and spiral GRAPPA are both compatible with real-time reconstruction. Reconstruction time on a single processor was 14.4 ms/frame for view sharing (1.8 ms per coil for nuFFT) and 116.4 ms/frame for spiral GRAPPA (102 ms for GRAPPA interpolation; 14.4 ms for nuFFT). Note that GRAPPA weights were precomputed (85 seconds). With eight processors and using parfor loops in MATLAB, our implementation of spiral GRAPPA produced a true reconstruction latency of 116 ms and kept up with the raw MRI data rate of 18 ms/frame. PI-CS with temporal finite difference constraint requires reconstructing all frames at once. Total reconstruction time for 5 seconds of RT-MRI data was 17.5 minutes. If view sharing and spiral GRAPPA were performed

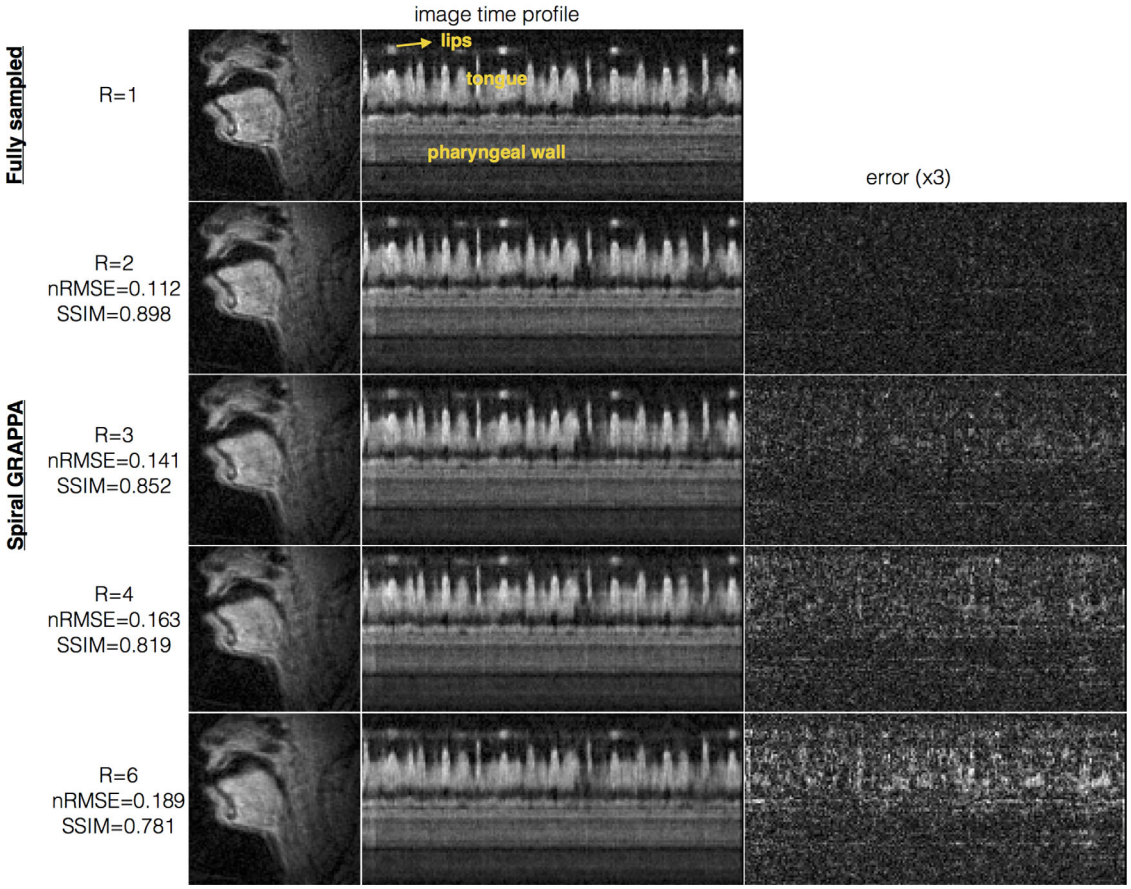

FIG. 3. Spiral GRAPPA reconstructions from retrospectively undersampled spiral data (12 interleaves/frame, 2.5 mm^2). The speaker repeated the phrase "ala-ara-asa-asha" at a normal pace. The first column shows an example spatial frame; the second column shows the image time profile marked by the solid white arrow; the third column shows the error time profiles, which are scaled up $\times 3$ for better visualization. Good spatiotemporal fidelity is maintained up to $R = 4$. At $R >= 6$, spatial blurring of the high-frequency edges is visually evident. The noise amplification with increasing R was consistent with the factor $\sqrt{N_{fs}/N_{us}}$; where N_{fs} and N_{us} are respectively the number of acquired interleaves in the fully sampled data and undersampled data.

serially frame by frame, they would have resulted in total reconstruction times of 4 and 32 seconds, respectively.

DISCUSSION

We have demonstrated the feasibility of accelerating RT-MRI of speech by up to 4-fold using spiral GRAPPA with reconstruction latency times of 116 ms in the MATLAB environment. This approach effectively leverages the advantages of a custom upper-airway phased-array coil and efficient spiral sampling. A time resolution of 18 ms/frame (55 fps) was achieved at a spatial resolution of 2.4 mm^2. Calibration time was approximately 12 seconds, which is negligible considering that a typical speech MRI data collection with a sequence of choice is 30 to 90 minutes. Good fidelity reconstructions were observed up to $R = 4$. At $R \geq 6$, spatial blurring at the edges was evident, which is attributed to the large gaps in k-space coupled with the limited number of receiver coils; note that the use of an array with a larger number of coil elements may increase the acceleration possible with through-time non-Cartesian GRAPPA, as has been shown in previous work (27–32).

The compatibility with pseudo-golden-angle interleaving allowed for retrospective adjustment of the time resolution to visualize rapidly moving articulators. Examples demonstrating its utility in visualizing rapid speech and beat boxing were presented. Attributable to the use of smaller temporal windows, superior temporal fidelity was observed with spiral GRAPPA and PI-CS in comparison to view sharing (3 interleaves/frame with spiral GRAPPA vs 13 interleaves/frame with view sharing). Reconstruction times with spiral GRAPPA were 8 times slower compared to view sharing. In addition, spiral GRAPPA involved a one-time computation of the GRAPPA weights (85 seconds). Parallel computing was exploited in spiral GRAPPA to achieve a true reconstruction latency of 116 ms, which kept up with the raw MRI data rate of 18 ms/frame when using eight processors. PI-CS required 17.5 minutes to jointly reconstruct 5 seconds of RT-MRI data. In comparison, a serial frame-by-frame reconstruction using view sharing or spiral GRAPPA would have resulted in total reconstruction times of 4 or 32 seconds, respectively.

The multishot spiral GRAPPA method in this work can further be extended in several ways. Performance may be improved through the use of higher channel

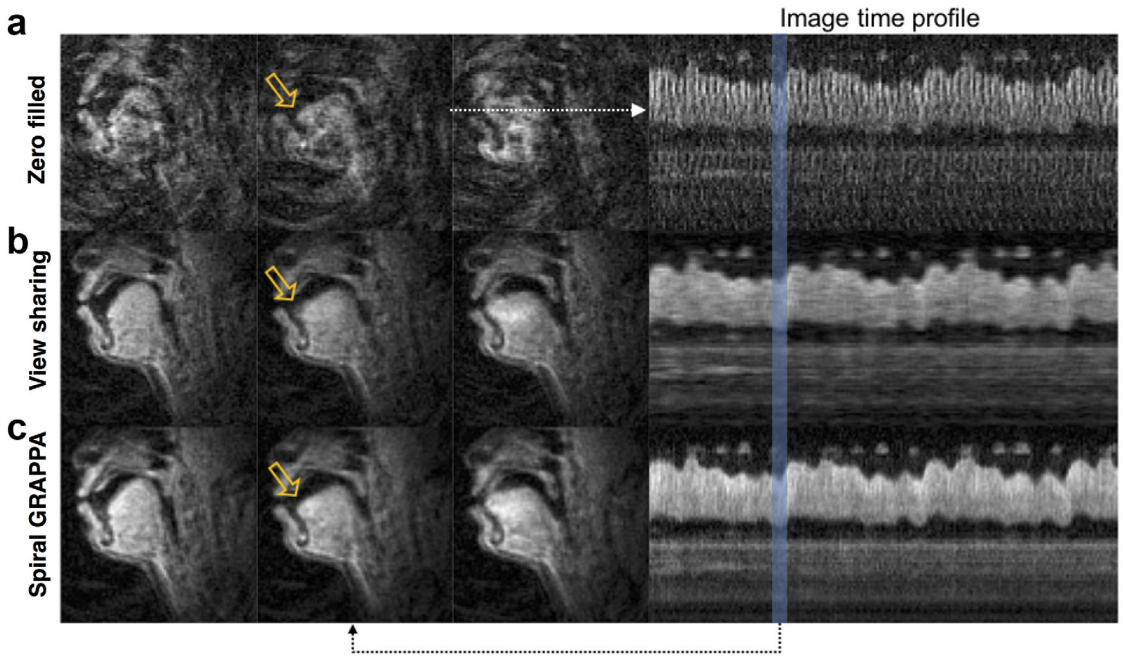

FIG. 4. RT-MRI of fluent speech with (**a**) zero-filled reconstruction at R = 4.3 (3 TR, 18 ms); (**b**) view-sharing reconstruction (step size: 3 TR; frame rate: 55 fps; native time resolution = 78 ms); (**c**) through-time GRAPPA reconstruction at R = 4.3 (3 TR, 18 ms). The subject produced the phrase "one-two-three-four-five" at a rapid pace. The first three columns show a sequence of spatial frames corresponding to the timing of the tongue tip hitting the palate, and the last column shows the image time profiles at the cross-section marked by the white dotted arrow. Zero-filled–based images show considerable undersampling artifacts. View sharing captured slow dynamic movements well, but blurred rapid movements. In contrast, through-time spiral GRAPPA provided improved temporal fidelity of rapidly moving articulators. For instance, subtle tongue tip shaping is seen to be captured with spiral GRAPPA while blurred with view sharing (see solid yellow arrows).

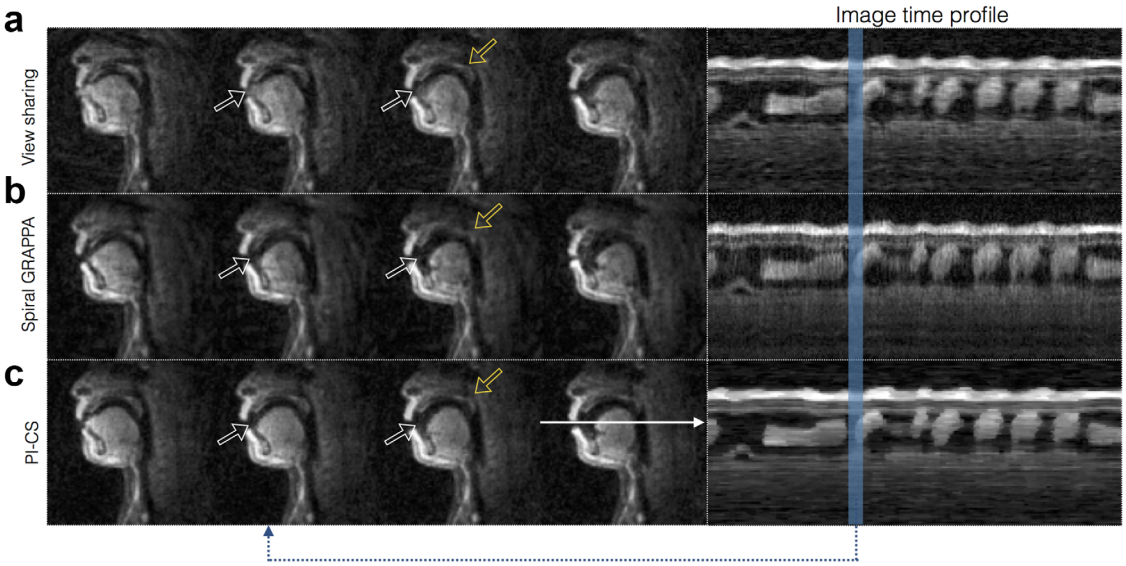

FIG. 5. RT-MRI of beat boxing at 55 fps with (**a**) view sharing, (**b**) spiral GRAPPA, and (**c**) PI-CS. The dynamic time series corresponded to the 5-second snippet of a free-style beat-boxing task. The last column shows the image time profile as marked by the solid white arrow. Columns 1 to 4 show the dynamic images of the time instance marked on the first column. In comparison to PI-CS, and spiral GRAPPA, view sharing resulted in temporal blurring of fast articulatory dynamics. For instance, the rapid inward movement of the tongue tip is captured with better fidelity in the spiral GRAPPA and PI-CS methods when compared to view sharing (see white arrows). Similarly, subtle movements of the velum are blurred with view sharing in comparison to spiral GRAPPA and PI-CS (see yellow arrows). PI-CS has improved signal to noise attributed to inherent denoising. Approximate MATLAB (The MathWorks, Inc., Natick, MA, USA) latency times were: view sharing (14 ms/frame); spiral GRAPPA (116 ms/frame); and PI-CS (17.5 mins to jointly reconstruct 5 secs of RT-MRI data).

count receiver arrays. Residual noise in the reconstructions may be mitigated by using on-the-fly denoising techniques, such as median filtering (45) or nonlocal means filtering (46). In this work, for simplicity, we have focused on assessing image quality as a function of the number of time frames in the calibration data with a fixed block size of 1×1 in the readout and azimuthal directions. There may be further room to reduce the 12-second calibration time by exploiting through k-space GRAPPA. We have demonstrated applicability of through-time non-Cartesian GRAPPA with spiral trajectories at 1.5T. However, it could be applied with radial trajectories in cases where spiral trajectories are challenged by off-resonance artifacts (e.g., at higher field strengths). Self-calibrated imaging (47) may be feasible when the number of interleaves in the fully sampled acquisition equals the desired acceleration level (four in this study).

Several applications of spiral GRAPPA-based RT-MRI of the upper airway remain to be explored. These include real-time feedback-based imaging during singer training, language training, and bidirectional speaker communication. Clinical applications requiring instant feedback may also benefit. These include speech rehabilitation applications such as in stuttering, aphasia, reorganization of speech post-treatment, or evaluation of swallowing difficulties (eg, dysphagia).

CONCLUSION

We have demonstrated the ability of spiral GRAPPA to accelerate RT-MRI of speech by up to 4-fold with a reconstruction latency of 116 ms. The approach provided single-slice RT-MRI of the upper airway at a time resolution of 18 ms/frame and spatial resolution of 2.4 mm^2/pixel. Attributable to the use of smaller temporal footprints per frame, this approach provides superior temporal fidelity in comparison to existing on-the-fly view-sharing schemes and comparable temporal fidelity to state-of-the-art PI-CS schemes that utilize temporal constraints. In contrast to PI-CS schemes, this approach can be perfectly parallelized and produces images frame by frame, making it applicable to a wide range of RT-MRI experiments of the upper airway that require real-time visualization of rapid articulatory movements.

REFERENCES

1. Bresch E, Kim Y-C, Nayak K, Byrd D, Narayanan S. Seeing speech: capturing vocal tract shaping using real-time magnetic resonance imaging. IEEE Signal Process Mag 2008;25:123–132.
2. Scott AD, Wylezinska M, Birch MJ, Miquel ME. Speech MRI: morphology and function. Phys Medica 2014;30:604–618.
3. Lingala SG, Sutton BP, Miquel ME, Nayak KS. Recommendations for real-time speech MRI. J Magn Reson Imaging 2016;43:28–44.
4. Atik B, Bekerecioglu M, Tan O, Etlik O, Davran R, Arslan H. Evaluation of dynamic magnetic resonance imaging in assessing velopharyngeal insufficiency during phonation. J Craniofac Surg 2008;19:566–572.
5. Bae Y, Kuehn DP, Conway CA, Sutton BP. Real-time magnetic resonance imaging of velopharyngeal activities with simultaneous speech recordings. Cleft Palate Craniofac J 2011;48:695–707.
6. Kazan-Tannus JF, Levine D, McKenzie C, Lim KH, Cohen B, Farrar N, Busse RF, Mulliken JB. Real-time magnetic resonance imaging aids prenatal diagnosis of isolated cleft palate. J Ultrasound Med 2005;24:1533–1540.
7. Perry JL, Kuehn DP, Wachtel JM, Bailey JS, Luginbuhl LL. Using magnetic resonance imaging for early assessment of submucous cleft palate: a case report. Cleft Palate Craniofac J 2012;49:e35–e41.
8. Zu Y, Narayanan SS, Kim YC, Nayak K, Bronson-Lowe C, Villegas B, Ouyoung M, Sinha UK. Evaluation of swallow function after tongue cancer treatment using real-time magnetic resonance imaging: a pilot study. JAMA Otolaryngol Head Neck Surg 2013;139:1312–1319.
9. Wu Z, Chen W, Khoo MC, Davidson Ward SL, Nayak KS. Evaluation of upper airway collapsibility using real-time MRI. J Magn Reson Imaging 2016;44:158-167.
10. 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;115:1771–1776.
11. Niebergall A, Zhang S, Kunay E, Keydana G, Job M, Uecker M, Frahm J. Real-time MRI of speaking at a resolution of 33 ms: undersampled radial FLASH with nonlinear inverse reconstruction. Magn Reson Med 2013;69:477–485.
12. Freitas AC, Wylezinska M, Birch MJ, Petersen S, Miquel ME. Comparison of Cartesian and non-Cartesian real-time MRI sequences at 1.5T to assess velar motion and velopharyngeal closure during speech. PLoS One 2016;11:e0153322.
13. Fu M, Zhao B, Carignan C, Shosted RK, Perry JL, Kuehn DP, Liang Z-P, Sutton BP. High-resolution dynamic speech imaging with joint low-rank and sparsity constraints. Magn Reson Med 2015;73:1820–1832.
14. Lingala SG, Zhu Y, Kim Y, Toutios A, Narayanan S, Nayak KS. A fast and flexible MRI system for the study of dynamic vocal tract shaping. Magn Reson Med 2017;77:112–125.
15. Burdumy M, Traser L, Richter B, Echternach M, Korvink JG, Hennig J, Zaitsev M. Acceleration of MRI of the vocal tract provides additional insight into articulator modifications. J Magn Reson Imaging 2015;42:925–935.
16. Iltis PW, Frahm J, Voit D, Joseph AA, Schoonderwaldt E, Altenmüller E. High-speed real-time magnetic resonance imaging of fast tongue movements in elite horn players. Quant Imaging Med Surg 2015;5:374–381.
17. Kerr AB, Pauly JM, Hu BS, Li KC, Hardy CJ, Meyer CH, Macovski A, Nishimura DG. Real-time interactive MRI on a conventional scanner. Magn Reson Med 1997;38:355–367.
18. Holsinger AE, Wright RC, Riederer SJ, Farzaneh F, Grimm RC, Maier JK. Real-time interactive magnetic resonance imaging. Magn Reson Med 1990;14:547–553.
19. Guttman M, Lederman R, Sorger J, McVeigh E. Real-time volume rendered MRI for interventional guidance. J Cardiovasc Magn Reson 2002;4:431–442.
20. Anon. Transmission Systems and Media, Digital Systems and Networks: Series G. 2003.
21. Claypool M, Claypool K. Latency and player actions in online games. Commun ACM 2006;49:40.
22. Fu M, Barlaz MS, Holtrop JL, Perry JL, Kuehn DP, Shosted RK, Liang ZP, Sutton BP. High-frame-rate full-vocal-tract 3D dynamic speech imaging. Magn Reson Med 2017;77:1619–1629.
23. Sorensen TS, Atkinson D, Schaeffter T, Hansen MS. Real-time reconstruction of sensitivity encoded radial magnetic resonance imaging using a graphics processing unit. Med Imaging IEEE Trans 2009;28:1974–1985.
24. Schaetz S, Uecker M. A multi-GPU programming library for real-time applications. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Vol. 7439 LNCS. 2012. pp. 114–128.
25. Uecker M, Zhang S, Frahm J. Nonlinear inverse reconstruction for real-time MRI of the human heart using undersampled radial FLASH. Magn Reson Med 2010;63:1456–1462.
26. Smith DS, Gore JC, Yankeelov TE, Welch EB. Real-time compressive sensing MRI reconstruction using GPU computing and split Bregman methods. Int J Biomed Imaging 2012;2012:864827.
27. Seiberlich N, Ehses P, Duerk J, Gilkeson R, Griswold M. Improved radial GRAPPA calibration for real-time free-breathing cardiac imaging. Magn Reson Med 2011;65:492–505.
28. Seiberlich N, Lee G, Ehses P, Duerk JL, Gilkeson R, Griswold M. Improved temporal resolution in cardiac imaging using through-time spiral GRAPPA. Magn Reson Med 2011;66:1682–1688.
29. Barkauskas KJ, Rajiah P, Ashwath R, Hamilton JI, Chen Y, Ma D, Wright KL, Gulani V, Griswold MA, Seiberlich N. Quantification of

left ventricular functional parameter values using 3D spiral bSSFP and through-time non-Cartesian GRAPPA. J Cardiovasc Magn Reson 2014;16:65.
30. Wright KL, Lee GR, Ehses P, Griswold MA, Gulani V, Seiberlich N. Three-dimensional through-time radial GRAPPA for renal MR angiography. J Magn Reson Imaging 2014;40:864–874.
31. Chen Y, Lee GR, Wright KL, Badve C, Nakamoto D, Yu A, Schluchter MD, Griswold MA, Seiberlich N, Gulani V. Free-breathing liver perfusion imaging using 3-dimensional through-time spiral generalized autocalibrating partially parallel acquisition acceleration. Invest Radiol 2015;50:367–375.
32. Chen Y, Lee GR, Aandal G, Badve C, WrighT KL, Griswold MA, Seiberlich N, Gulani V. Rapid volumetric T1 mapping of the abdomen using three-dimensional through-time spiral GRAPPA. Magn Reson Med 2016;75:1457–1465.
33. Kim YC, Narayanan SS, Nayak KS. Flexible retrospective selection of temporal resolution in real-time speech MRI using a golden-ratio spiral view order. Magn Reson Med 2011;65:1365–1371.
34. Santos JM, Wright GA, Pauly JM. Flexible real-time magnetic resonance imaging framework. Conf Proc IEEE Eng Med Biol Soc 2004;2: 1048–1051.
35. Fessler JA, Sutton BP. Nonuniform fast Fourier transforms using min-max interpolation. IEEE Trans Signal Process 2003;51:560–574.
36. Walsh DO, Gmitro AF, Marcellin MW. Adaptive reconstruction of phased array {MR} imagery. Magn Reson Med 2000;43:682–690.
37. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 2004;13:600–612.
38. Doneva M, Stehning C, Nehrke K, Börnert P. Improving scan efficiency of respiratory gated imaging using compressed sensing with 3D Cartesian golden angle sampling. In Proceedings of the 19th Annual Meeting of ISMRM, Montreal, Canada, 2011. p. 641.
39. Kim YC, Lebel RM, Wu Z, Ward SLD, Khoo MCK, Nayak KS. Real-time 3D magnetic resonance imaging of the pharyngeal airway in sleep apnea. Magn Reson Med 2014;71:1501–1510.
40. Han X, Wright K, Gulani V, Seiberlich N. Golden angle through-time radial GRAPPA for real-time cardiac MRI. In Procedings of the 21st Annual Meeting of ISMRM, Salt Lake City, Utah, USA, 2013. p. 3834.
41. Proctor M, Bresch E, Byrd D, Nayak K, Narayanan S. Paralinguistic mechanisms of production in human "beatboxing": a real-time magnetic resonance imaging study. J Acoust Soc Am 2013;133:1043–1054.
42. Bresch E, Nielsen J, Nayak K, Narayanan S. Synchronized and noise-robust audio recordings during realtime magnetic resonance imaging scans. J Acoust Soc Am 2006;120:1791–1794.
43. Vaz C, Ramanarayanan V, Narayanan S. A two-step technique for MRI audio enhancement using dictionary learning and wavelet packet analysis. In: Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH. 2013. pp. 1312–1315.
44. Frahm J, Schaetz S, Markus U, Zhang S, Voit D, Merboldt KD, Sohns J, Lotz J, Uecker M. On the temporal fidelity of nonlinear inverse reconstructions for real-time MRI-the motion challenge. Open Med Imaging J 2014;8:1–7.
45. Ataman E, Aatre VK, Wong KM. A fast method for real-time median filtering. IEEE Trans Acoust 1980;28:415–421.
46. Klosowski J, Frahm J. Image denoising for real-time MRI. Magn Reson Med 2017;77:1340–1352.
47. Hamilton JI, Wright KL, Griswold MA, Seiberlich N. Self-calibrating interleaved reconstruction for through-time non-Cartesian GRAPPA. In Procedings of the 21st Annual Meeting of ISMRM, Salt Lake City, Utah, USA, 2013. p. 3836.

SUPPORTING INFORMATION

Additional supporting information can be found in the online version of this article.

Video S1. Spiral GRAPPA reconstructions from retrospectively undersampled spiral data (12 interleaves/frame, 2.5 mm^2). The speaker repeated the phrase "ala-ara-asa-asha" at a normal pace. The reconstructions are shown for subsampling factors ($R = 2, 3, 4, 6$). The error images are scaled up by a factor of 3 for better visualization.

Video S2. RT-MRI of beat boxing at 55 frames per second with spiral GRAPPA reconstruction and noise-cancelled audio.