Shortening total acquisition times for high-speed dynamic speech magnetic resonance imaging using temporally sparse navigation

Riwei Jin, Peng Xi, Brad Sutton

Department of Bioengineering, University of Illinois Urbana-Champaign
Beckman Institute for Advanced Science and Technology

Introduction

- **Dynamic magnetic resonance imaging** (MRI) has seen increasing use in the speech imaging field:
  - Great soft tissue contrast for linguistics, language
  - Capturing articulatory changes
  - Drawback has been relatively slow imaging speed
- Recent methods utilizing the **Partial Separability** (PS) model[1,2]
  - Temporal resolution up to 166 fps
  - But long acquisition time up to 10 minutes
- We propose **sparsely sampling** the temporal navigator and associating each navigator acquisition with multiple k-space lines.
  - Decrease the temporal resolution to about 40 fps, which is still sufficient for articulatory motions.
  - But also decrease the acquisition time by 40%

Methods

- Partial Separability (PS) model
  - The PS model assumes strong spatiotemporal correlation exists.
  - We can express our desired spatiotemporal image $f(\mathbf{r},t)$as the summation of product of low rank constrained temporal navigator bases and image bases through PS model:

$$f(\mathbf{r},t) = \sum_{l=1}^{L} \Psi_l(\mathbf{r})\phi_l(t)$$

  - L is the model order, usually we use 10 at a 2D case because of Fig.1. $\{\Psi_l(\mathbf{r})\}_{l=1}^L$ denotes a set of spatial basis functions, and $\{\phi_l(t)\}_{l=1}^L$ denotes a set of temporal basis functions.

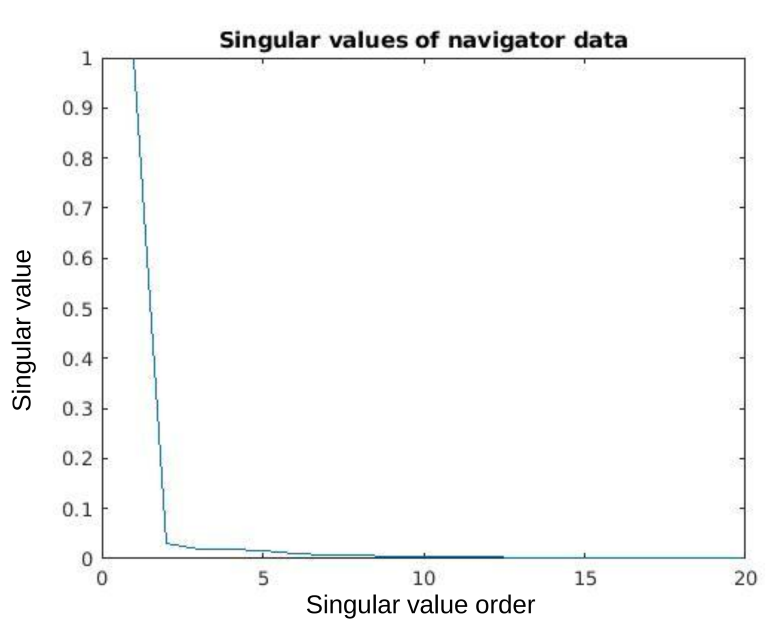


Fig.1 Singular values of navigator data. It drops quickly at first 5 orders; in this case we can say first 10 orders mostly covers the dynamics

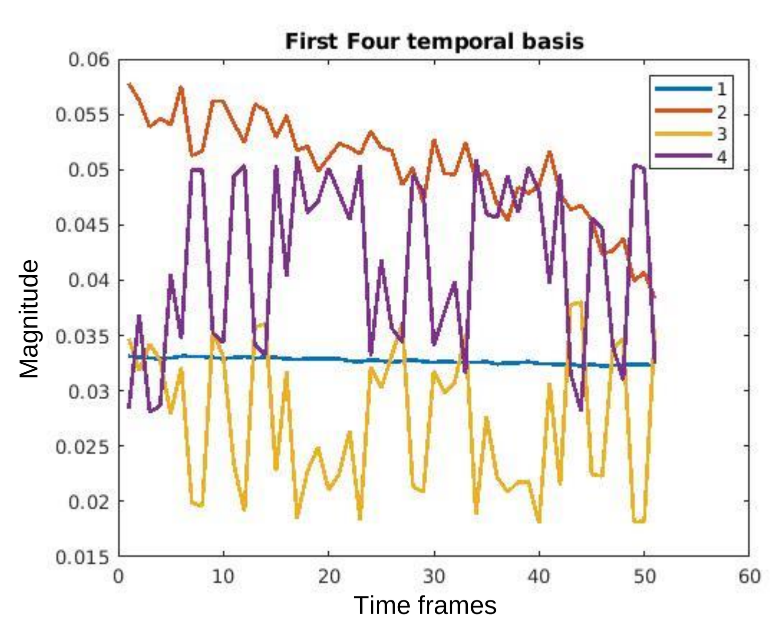


Fig.2 Magnitude plot of first four temporal basis of 50 time-frames.

Sparse Sampling

We sparsely sample two sets of (k, t)-space data in an interlaced manner: an imaging dataset (from which we will estimate $\{\Psi_l(\mathbf{r})\}_{l=1}^L$) and a temporal navigator dataset (which provides estimates of $\{\phi_l(t)\}_{l=1}^L$).[1] For the navigator dataset, we used a spiral-trajectory to sample k-space. For the imaging dataset, a Cartesian trajectory with random phase encoding to acquire imaging data with high spatial resolution and structural image quality.

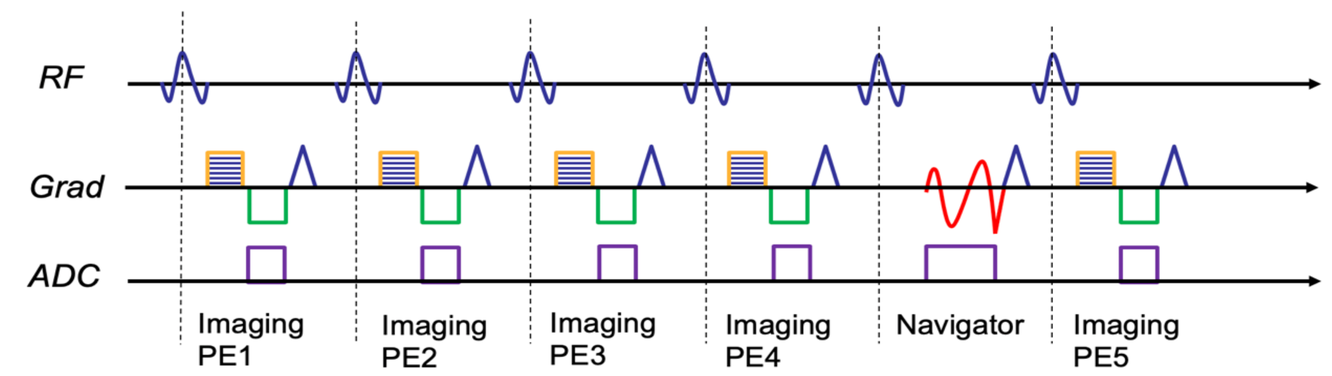


Fig.3 shows a simplified pulse sequence diagram for the MRI sequence that samples random ky lines of Imaging datasets and uses a spiral navigator.

Reconstruction

- For temporal basis, we do Singular Value Decomposition
- For spatial basis, we do a Least Square problem with Huber Penalty

$$\{\widehat{\Psi_l}(\mathbf{r})\}_{l=1}^L = \underset{\{\Psi_l(\mathbf{r})\}_{l=1}^L}{\operatorname{argmin}} \left\{ \left\| s(\boldsymbol{k},t) - E\left[\sum_{l=1}^{L} \Psi_l(\mathbf{r})\varphi_l(t)\right] \right\|_2^2 + \lambda\phi_{huber}[\Psi_l(\mathbf{r})] \right\}$$

where $\phi_{huber}(u) = \begin{cases} u^2 & |u| \le M \\ M(2|u| - M) & |u| > M \end{cases}$

- $s(\boldsymbol{k},t)$ represents the original data signal, E represents a system calculating matrix. M is the Huber penalty parameter
- Apply PS model to get back whole timeseries: $\hat{f}(\mathbf{r},t) = \sum_{l=1}^{L} \widehat{\Psi_l}(\mathbf{r})\varphi_l(t)$

Results

- We acquire 256-matrix size, 24 cm field of view, 6-mm mid-sagittal slice, Spatial resolution: 0.94×0.94×6 mm.
- We reduce the temporal resolution of our model by 4 which provides 43 images per second instead of 108. This frame rate is still high enough to visualize the articulatory motions during speech.
- Total duration of the scan is reduced to 62.5% In this case, we acquired 45 full frames of data in 66 sec (instead of 105 sec), resulting in 2880 reconstructed images from the PS model at 43 frames per second.
- Figure 4 and 5 shows the results from the 2D single-slice experiment where the subject counted to four.

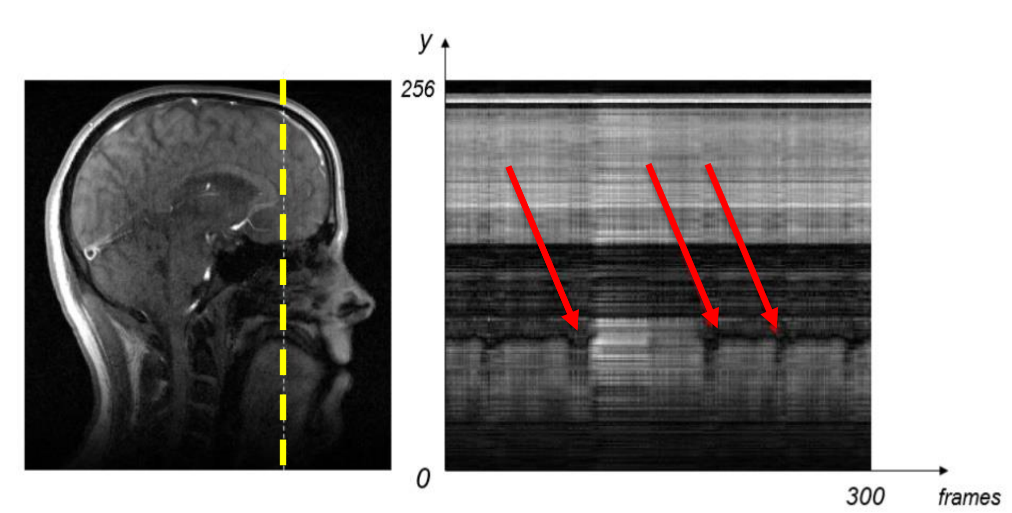


Fig.4 shows the temporal profile taken along a vertical strip across the tongue which displays tongue movement for one utterance of 'one two three four' in 300 time frames. The red arrows show the vertical movement of tongue

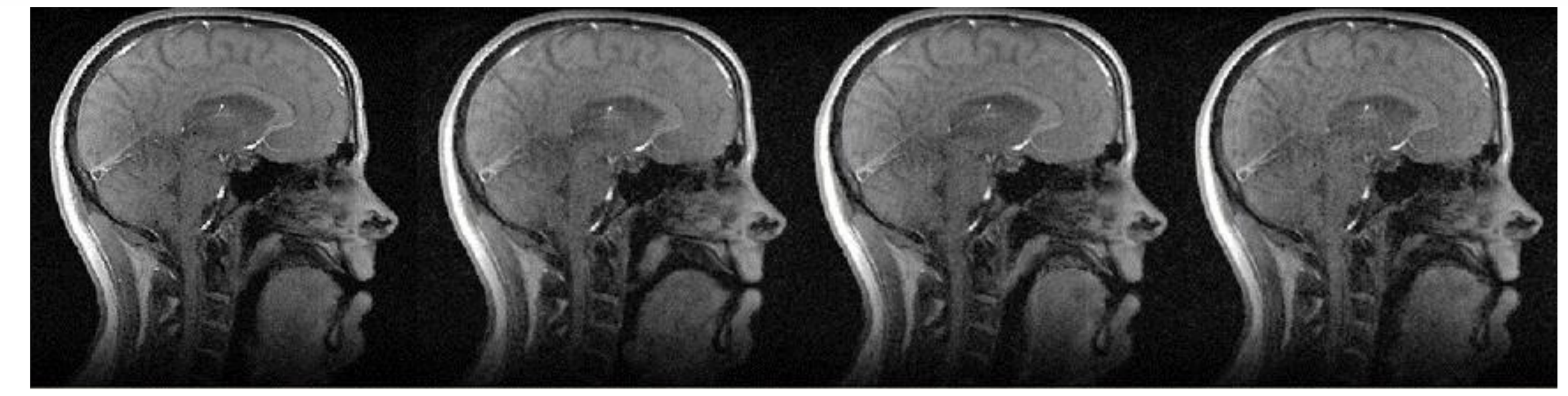


Fig.5 Articulatory and tongue motions at four different time instances

Extend to 3D

Sequence diagram

We add another slice encoding gradient in z-direction which performs a random phase encoding in both y and z direction in figure 6. All other parts are similar to 2D sequence

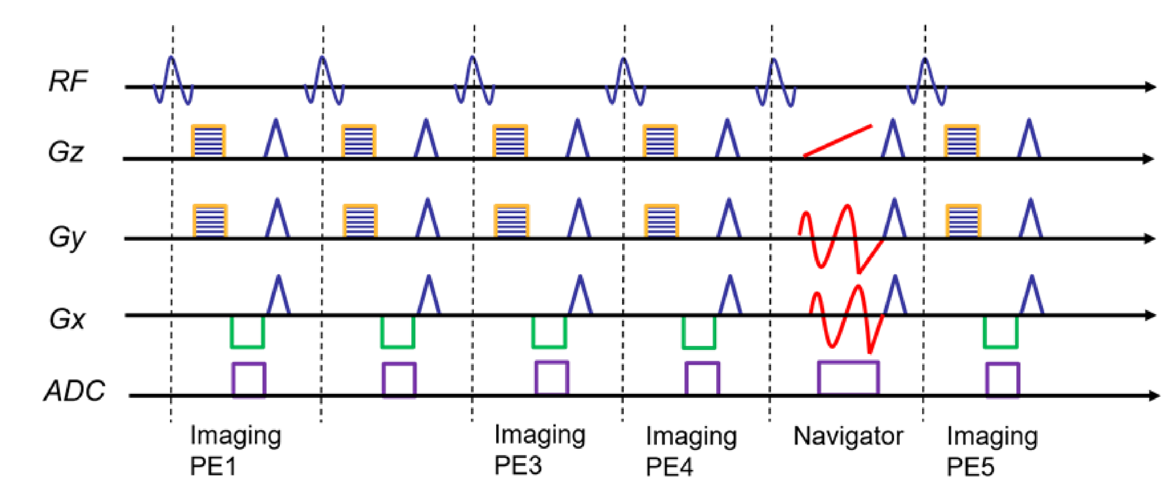
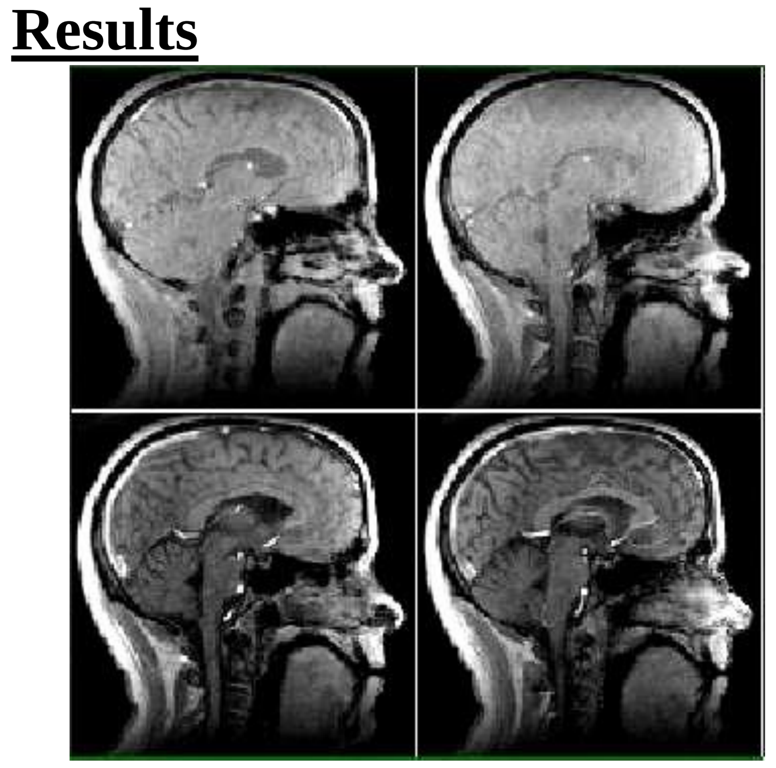


Fig.6 shows a simplified pulse sequence diagram for the 3D MRI sequence that samples random ky and kz lines of Imaging datasets and uses a spiral navigator.

Results



- We acquire 4 by 128-matrix size, 6-mm slice thickness. Spatial resolution: 0.94×0.94×6 mm.
- 5 minutes acquisition time with 40 fps
- Model order increase to 30 to cover more dynamics in 3D.

Fig.6 A movie of 4 by 128 matrix size , 24 cm field of view, 6-mm slice thickness. Model order 30. Subject counted from one to four.

Works Referenced

1. Liang Z-P. Spatiotemporal imaging with partially separable functions. In Proceedings of IEEE International Symposium on Biomedical Imaging, Washington D.C., USA, 2007. pp. 988–991.
2. Fu, M. , Zhao, B. , Carignan, C. , Shosted, R. K., Perry, J. L., Kuehn, D. P., Liang, Z. and Sutton, B. P. (2015), High‐resolution dynamic speech imaging with joint low‐rank and sparsity constraints. Magn. Reson. Med., 73: 1820-1832. doi:10.1002/mrm.25302
3. Fu, M. , Barlaz, M. S., Holtrop, J. L., Perry, J. L., Kuehn, D. P., Shosted, R. K., Liang, Z. and Sutton, B. P. (2017), High‐frame‐rate full‐vocal‐tract 3D dynamic speech imaging. Magn. Reson. Med., 77: 1619-1629. doi:10.1002/mrm.26248