High Temporal Resolution Motion Estimation Using a Self-Navigated Simultaneous Multi-Slice Echo Planar Imaging Acquisition

Jose R. Teruel, PhD^{1,2}, Joshua M. Kuperman, PhD¹, Anders M. Dale, PhD^{1,3}, and Nathan S. White, PhD^{1,*}

¹Department of Radiology, University of California San Diego, La Jolla, California, USA

²Department of Radiation Oncology, NYU Langone Health, New York, New York, USA

³Department of Neurosciences, University of California San Diego, La Jolla, California, USA

Abstract

Background—Subject motion is known to produce spurious covariance among time-series in functional connectivity that has been reported to induce distance-dependent spurious correlations.

Purpose—To present a feasibility study for applying the extended Kalman filter (EKF) framework for high temporal resolution motion correction of resting state functional MRI (rs-fMRI) series using each simultaneous multi-slice (SMS) echo planar imaging (EPI) shot as its own navigator.

Study Type—Prospective feasibility study.

Population/Subjects—Three human volunteers.

Field Strength/Sequence—3T GE DISCOVERY MR750 scanner using a 32-channel head coil. Simultaneous multi-slice rs- fMRI sequence with repetition time (TR)/echo time (TE) = 800/30 ms, and SMS factor 6.

Assessment—Motion estimates were computed using two techniques: a conventional rigid-body volume-wise registration; and a high-temporal resolution motion estimation rigid-body approach. The reference image was resampled using the estimates obtained from both approaches and the difference between these predicted volumes and the original moving series was summarized using the normalized mean squared error (NMSE).

Statistical Tests—Direct comparison of NMSE values.

Results—High-temporal motion estimation was always superior to volume-wise motion estimation for the sample presented. For staged continuous rotations, the NMSE using high-temporal resolution motion estimates ranged between [0.130, 0.150] for the first volunteer (in-

*Address reprint requests to: N.S.W., Department of Radiology, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0515. nswhite@ucsd.edu.

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

Level of Evidence: 2

Technical Efficacy: Stage 1

plane rotations), between [0.060, 0.068] for the second volunteer (in-plane rotations), and between [0.063, 0.080] for the third volunteer (through-plane rotations). These values went up to [0.384, 0.464]; [0.136, 0.179]; and [0.080, 0.096], respectively, when using volume-wise motion estimates.

Data Conclusion—Accurate high-temporal rigid-body motion estimates can be obtained for rs-fMRI taking advantage of simultaneous multi-slice EPI sub-TR shots.

Subject motion is recognized as one of the main sources of functional connectivity uncertainties for resting state functional MRI (fMRI).¹ Subject motion is known to produce spurious covariance among time-series in functional connectivity that has been reported to induce distance-dependent spurious correlations.^{2,3} Issues related to subject motion are particularly severe for populations that tend to exhibit continuous motion across time such as children and the elderly. For instance, spurious effects in developmental resting state fMRI due to residual motion have been reported in resting state fMRI studies for children and the elder population.^{4,5}

The most common and widespread approach to manage subject motion in resting state fMRI is to perform retrospective volume-wise rigid realignment for each series acquired. While this technique can help mitigate some motion related problems, it cannot resolve intra-volume (sub-repetition time [TR]) motion, i.e., different slices within the same volume might be affected by different magnitudes of motion. This might produce unreliable estimates of motion particularly for cases where substantial sub-TR motion is present. In addition, lower precision in motion estimates may affect motion summary metrics such as framewise displacement (FD), that are used to censor frames corrupted by sudden motion (one volume to the next).^{2,3,6}

The fMRI is conventionally acquired slice-by-slice in 2D echo planar imaging (EPI). However, the recent development of simultaneous multi-slice (SMS) provides a tool to obtain several slices at once (one shot) that are prospectively separated into individual slices.^{7–12} SMS is used to decrease the acquisition time for each fMRI run, while at the same time provides a subset (depending on the multislice factor) of the complete 3D volume, where all slices are obtained at the same time, i.e., no motion among slices acquired simultaneously. Therefore, this subset of 3D volumetric information can be used to produce high temporal motion estimates for each shot.

The purpose of the current work is to present an initial feasibility study for applying a previously described image based tracking method¹³ for higher temporal resolution motion estimation using each multi-slice EPI shot as its own navigator (i.e., self-navigated).

Materials and Methods

The study was approved by the local Institutional Review Board. Three healthy volunteers (one male, 33 years old; two females, 24 and 26 years old) were recruited for the study and provided written informed consent.

Image Acquisition

For each volunteer six resting-state fMRI (rs-fMRI) runs were acquired on a 3T GE DISCOVERY MR750 scanner using a 32-channel head coil with the following parameters: TR/TE (ms) = 800/30; in plane resolution = $2.4 \times 2.4 \text{ mm}^2$; matrix = 90×90 ; number of slices = 60; slice thickness = 2.4 mm; simultaneous multi-slice factor = 6; number of temporal frames = 380; and echo planar imaging readout; axial in-plane orientation. The scanning time to obtain the six rs-fMRI runs was slightly over 30 min. Each SMS acquisition was reconstructed offline using the raw data. Two of the volunteers were asked to remain as still as possible for three of the rs-fMRI runs, while for the other three runs, they were asked to rotate their head continuously around the z axis (yaw rotation). The third volunteer was instructed to remain still for two of the rs-fMRI runs, to rotate the head around the x-axis (pitch rotation or nodding) for two runs, and to displace the head to different positions within the coil and remain in that location (sudden motion) for the last two runs. For each run where the volunteers were asked to move continuously (z or x rotations continuously through the run) the instruction to start moving was given approximately 30 seconds into the run, and they were also instructed to rest for the last 30 seconds of the run. These motion experiments were designed to assess the accuracy and precision of both through-plane and in-plane subject motion.

Motion Estimation

The average volume of the first rs-fMRI run without staged motion was used as the reference volume for motion estimation for each volunteer. Rigid-body motion estimates were obtained in two different ways for all staged-motion fMRI runs: Conventional volume-wise rigid motion estimation using mcflirt (FSL, FMRIB, Oxford, UK)¹⁴; and the proposed framework with high resolution motion estimation using each simultaneous multi-slice shot as its own navigator (i.e., self-navigated) as described below.

Self-navigated Motion Estimation

Three-dimensional (3D) rigid body motion estimation was carried out using the extended Kalman filter (EKF) algorithm as applied previously by White et al¹³ to obtain real-time motion tracking in structural brain imaging using navigators. The EKF algorithm provides recursive motion estimates in nonlinear dynamic systems perturbed by Gaussian noise.¹⁵ For the purpose of this study, the EKF algorithm was applied retrospectively to the data after all the imaging frames were collected. In addition, to prevent areas like the neck and jaw from contributing adversely to the motion estimates, a brain region of interest (ROI) was obtained using a T_2^* -EPI brain atlas registered to the reference image. Only the subset of voxels included in the ROI were used for motion estimation.

Image Resampling

To establish the performance of the motion estimates, the reference volume was resampled to the original moving image using the produced motion estimates. For the volume-wise motion estimation approach, one set of rigid motion estimates was obtained for each frame and used to resample the reference volume to native (moving) space. For the high-temporal resolution motion estimation approach, one set of rigid motion estimates was obtained for

each multi-slice shot. Therefore, the 3D reference volume was resampled for each of the multi-shot motion estimates. After obtaining the high-temporal resolution resampled reference image, the slice selection order of the simultaneous multi-slice algorithm was applied to form a final 4D spatiotemporal image with the same temporal resolution as the original moving image.

The reference volumes resampled to the moving image space will be referred from now on as predicted volumes as they represent the location of the reference image predicted using the motion estimates for each approach.

Motion Estimates Accuracy

The accuracy of the motion estimation approaches presented can be evaluated by comparing the predicted time-series to the original time-series using the mean squared error normalized by the square of the mean or the original image volume:

$$\text{Normalized mean squared error (NMSE)} = \frac{1}{n} \sum_{i=1}^n \frac{(x_i - y_i)^2}{(\bar{x})^2}$$

where n is the number of voxels for each 3D volume, x_i is the intensity value for voxel i of the original moving image, $\bar{x}$ is the mean signal intensity of the 3D original moving volume and y_i is the intensity value for voxel i of the predicted volume. To summarize the NMSE for each 4D spatiotemporal run, we calculated the mean and standard deviation of the NMSE across each time-series.

Results

In Figure 1A, it is shown how the proposed correction framework successfully estimated sub-TR staged rotations of up to 10 degrees for one of the series where the volunteer was prompted to rotate their head around the z axis (in-plane yaw rotation). In Figure 1B, the rotation estimates for the same temporal volumes obtained using volume-wise motion correction are presented. By comparing Figure 1A and Figure 1B, it can be observed how sub-TR motion cannot be detected using volume-wise motion correction. Particularly, the volume-wise motion estimates resemble an approximated average of the previous high-resolution estimates that might be very variable for fast paced motion. Equivalently, Figure 1C and 1D depict the high resolution and volume-wise rotation estimates for one subject prompted to rotate their head around the x-axis (through-plane roll rotation). The motion estimates using volume-wise estimation resemble an approximate average of sub-TR motion resulting in underestimated and delayed values of rotations. The NMSE metric between the volumes of the original image and the predicted images are reported in Figure 2 for one run with continuous staged motion. As presented, there is a large difference between predicted volumes resampled using volume-wise motion estimates and the original moving image.

However, there is a lower difference between the predicted volumes resampled using super resolution motion estimates and the original moving image. To summarize the differences between predicted volumes and the original moving image, the mean and standard deviation

across all temporal frames for each series with staged motion and at rest are reported in Table 1. Our results outline how the difference between the predicted volume using volume-wise motion estimates and the original moving image is always higher than the difference between the predicted volume using high temporal resolution motion estimates and the original image using the described ID metric.

For staged continuous rotations, the normalized mean squared error (NMSE) using high temporal resolution motion estimates ranged between [0.130, 0.150] for the first volunteer (in-plane rotations), between [0.060, 0.068] for the second volunteer (in-plane rotations), and between [0.063, 0.080] for the third volunteer (through-plane rotations). These values went up to [0.384, 0.464]; [0.136, 0.179], and [0.080, 0.096], respectively, when using volume-wise motion estimates. For motion-free runs, the NMSE for each series were found to be very low compared with the runs with motion and very similar for both estimation methods as might be expected. In Figure 3, an example of the subtracted image between predicted volumes and the original image in one series acquired with staged continuous motion along the z-axis is presented. To provide a complete depiction of the difference between predicted volumes and the original moving volume, Supporting Video S1, which is available online, has been included as supplemental material presenting the “moving” middle slice for one complete series with staged continuous in-plane motion. This animation is an extension of Figure 3 for all frames acquired.

For the two series including sudden fixed displacements, the sharp displacements were promptly detected using high temporal resolution estimates, while for volume-wise estimates the displacement was not correctly depicted until a complete volume was acquired. This effect is presented in Figure 4 where we can observe the evolution of the NMSE metric around a sudden head displacement (transition motion). In Figure 4, additional panels have been included to present the original, predicted and subtracted images at each step presented in the NMSE plot, i.e., before, during, and after the transition motion occurred.

Discussion

Our results demonstrate how it is possible to use simultaneous multi-slice acquisitions to obtain accurate high temporal (sub-TR) motion estimates using each individual simultaneous multi-slice shot as a navigator. In this study, we have used a robust motion estimation framework described by White et al¹³ that is commonly used as a scanner product real-time tracking motion approach, and can be used as well for retrospective motion estimation offline as we have presented.

The importance of motion detection and correction for rs-fMRI analysis is very well known and described in the literature.¹ However, a majority of previous studies are mostly focused on how to improve rs-fMRI analysis in the presence of motion after volume-wise correction.^{3,16–19} For instance, motion regression using rigid estimates has been explored using the six motion estimates and its first derivative for two to three timepoints (24 – 36 parameters in total). Yet, analysis including up to three timepoints of motion estimates was found to leave significant motion related variance in the data.^{16–18}

Another approach to avoid the nuisance effects of motion in rs-fMRI is the use of censoring. Censoring of data works by excluding frames that are highly affected by relative motion by means of a relative motion metric like FD.^{3,16–19} While this technique effectively removes the more affected frames, there are some limitations to it. For instance, the optimal threshold to censor a volume is not yet defined. In addition, there are unaddressed issues related to perform group statistics including datasets that were collected with the same number of frames but have been censored to different extents.

In this study, we explored a different approach to improve motion estimation for temporal acquisitions that focused on obtaining higher temporal resolution estimates of motion using SMS-EPI acquisitions.²⁰ Our results show how we can track the motion at sub-TR intervals using each SMS shot as a navigator. For our study with a TR of 800 ms, 60 slices and SMS factor 6, we collect 10 SMS shots per TR, resulting in a temporal resolution for motion estimation of 80 ms or 12.5 Hz. Because each SMS shot is only a subset of the complete 3D volume, we use the motion estimates to resample the reference volume to the original moving space (extending the reference 3D volume to the appropriate number of temporal frames for each motion estimation approach). Our results clearly show how the similarity between the predicted volumes and the original moving ones is much higher using the high-temporal motion estimates for resampling. This higher similarity applied to all presented scenarios in this study (staged continuous in-plane motion, staged continuous through-plane motion and during transition motion within the head coil). It is worth mention that it is well-recognized that motion. In addition, to optimize our approach we used an approximate brain mask for motion estimation that it is not generally used in conventional rigid motion registration approaches. Furthermore, the discussion for this technical development is limited to the scope of producing the high temporal motion estimates. However, its application to produce complete motion corrected images is currently being investigated and tested with promising results.²²

It should be acknowledged that this study uses an SMS factor of 6 motivated by its use in undergoing studies in our institution. While our motion estimation approach shows promising results using this SMS factor, the use of this algorithm with different SMS factors, particularly smaller ones, should be carefully evaluated. Finally, it is worth noting that the focus of this technical development is to present the feasibility of sub-TR motion estimation using a previously developed motion estimation framework, and, therefore, the detailed description of the framework has been omitted. Nevertheless, the algorithm is described in detail in a previous publication.¹³ Regarding study design limitations, for this feasibility study we had to design our cohort to only include volunteers that were able to introduce fast paced continuous motion during an extended period of time. Therefore, we limited our volunteers to young healthy adults that previously volunteered for brain MR imaging, reducing the number of available subjects.

In summary, we have demonstrated that accurate high-temporal rigid-body motion estimates can be obtained for rs-fMRI taking advantage of simultaneous multi-slice EPI sub-TR shots. This development may provide a tool to mitigate the effects of motion in fMRI data, particularly for subject populations inherently prone to motion as children and is expected to have a positive impact in studies targeting these populations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Contract grant sponsor: National Institutes of Health; contract grant number: U24DA041123; Contract grant sponsor: General Electric; contract grant number: Investigator Initiated Research, Award BOK92322

References

1. Power JD, Schlaggar BL, Petersen SE. Recent progress and outstanding issues in motion correction in resting state fMRI. *Neuroimage*. 2015; 105:536–551. [PubMed: 25462692]
2. Van Dijk KRA, Sabuncu MR, Buckner RL. The influence of head motion on intrinsic functional connectivity MRI. *Neuroimage*. 2012; 59:431–438. [PubMed: 21810475]
3. Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage*. 2012; 59:2142–2154. [PubMed: 22019881]
4. Fair DA, Dosenbach NUF, Church JA, et al. Development of distinct control networks through segregation and integration. *Proc Natl Acad Sci U S A*. 2007; 104:13507–13512. [PubMed: 17679691]
5. Andrews-Hanna JR, Snyder AZ, Vincent JL, et al. Disruption of large-scale brain systems in advanced aging. *Neuron*. 2007; 56:924–935. [PubMed: 18054866]
6. Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*. 2002; 17:825–841. [PubMed: 12377157]
7. Larkman DJ, Hajnal JV, Herlihy AH, Coutts GA, Young IR, Ehnholm G. Use of multicoil arrays for separation of signal from multiple slices simultaneously excited. *J Magn Reson Imaging*. 2001; 13:313–317. [PubMed: 11169840]
8. Breuer FA, Blaimer M, Heidemann RM, Mueller MF, Griswold MA, Jakob PM. Controlled aliasing in parallel imaging results in higher acceleration (CAIPIRINHA) for multi-slice imaging. *Magn Reson Med*. 2005; 53:684–691. [PubMed: 15723404]
9. Moeller S, Yacoub E, Olman CA, et al. Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magn Reson Med*. 2010; 63:1144–1153. [PubMed: 20432285]
10. Feinberg DA, Moeller S, Smith SM, et al. Multiplexed echo planar imaging for sub-second whole brain FMRI and fast diffusion imaging. *PLoS One*. 2010; 5:e15710.doi: 10.1371/journal.pone.0015710 [PubMed: 21187930]
11. Setsompop K, Gagoski BA, Polimeni JR, Witzel T, Wedeen VJ, Wald LL. Blipped-controlled aliasing in parallel imaging for simultaneous multislice echo planar imaging with reduced g-factor penalty. *Magn Reson Med*. 2012; 67:1210–1224. [PubMed: 21858868]
12. Feinberg DA, Setsompop K. Ultra-fast MRI of the human brain with simultaneous multi-slice imaging. *J Magn Reson*. 2013; 229:90–100. [PubMed: 23473893]
13. White N, Roddey C, Shankaranarayanan A, et al. PROMO: real-time prospective motion correction in MRI using image-based tracking. *Magn Reson Med*. 2010; 63:91–105. [PubMed: 20027635]
14. Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. FSL. *Neuroimage*. 2012; 62:782–790. [PubMed: 21979382]
15. Kalman RE. A new approach to linear filtering and prediction problems. *J Basic Eng*. 1960; 82:35–45.
16. Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL, Petersen SE. Methods to detect, characterize, and remove motion artifact in resting state fMRI. *Neuroimage*. 2014; 84:320–341. [PubMed: 23994314]

17. Satterthwaite TD, Elliott MA, Gerraty RT, et al. An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *Neuroimage*. 2013; 64:240–256. [PubMed: 22926292]
18. Yan C-G, Cheung B, Kelly C, et al. A comprehensive assessment of regional variation in the impact of head micromovements on functional connectomics. *Neuroimage*. 2013; 76:183–201. [PubMed: 23499792]
19. Yan C-G, Craddock RC, He Y, Milham M. Addressing head motion dependencies for small-world topologies in functional connectomics. *Front Hum Neurosci*. 2013; 7:910. [PubMed: 24421764]
20. Barth M, Breuer F, Koopmans PJ, Norris DG, Poser BA. Simultaneous multislice (SMS) imaging techniques. *Magn Reson Med*. 2016; 75:63–81. [PubMed: 26308571]
21. Beall EB, Lowe MJ. SimPACE: generating simulated motion corrupted {BOLD} data with synthetic-navigated acquisition for the development and evaluation of SLOMOCO: a new, highly effective slice-wise motion correction. *Neuroimage*. 2014; 101:21–34. [PubMed: 24969568]
22. Teruel JR, White NS, Brown TT, Kuperman JM, Dale AM. Super Resolution Motion Correction (SUPREMO) using simultaneous multi-slice EPI based fMRI; Proceedings of the 25th Annual Meeting of ISMRM; Honolulu. 2017. (abstract 3837)

FIGURE 1.

Detail of rigid motions estimates for two staged continuous motion runs. **A:** Rotations obtained using high temporal resolution motion estimation for an in-plane staged continuous motion run (Volunteer 1, staged continuous motion run 3). **B:** Rotations obtained using volume-wise motion estimation for an in-plane staged continuous motion run (Volunteer 1, staged continuous motion run 3). **C:** Rotations obtained using high temporal resolution motion estimation for a through-plane staged continuous motion run (Volunteer 3, staged continuous motion run 1). **D:** Rotations obtained using volume-wise motion estimation for a through-plane staged continuous motion run (Volunteer 3, staged continuous motion run 1).

FIGURE 2.

Plot of the normalized mean squared error (NMSE) metric between the original moving image and predicted volume resampled using high-temporal motion estimates (blue line) and volume-wise motion estimates (red line) (Volunteer 1, staged continuous motion run 3). The black broken line indicates temporal frame when the volunteer started rotating the head. The black solid box indicates the frame range that corresponds to the motion estimates depicted in Figure 1A and 1B.

FIGURE 3.

A: Slice from original moving volume (Volunteer 1, staged continuous motion run 3). **B:** Predicted volume obtained by resampling the reference volume using the high-temporal motion estimates. **C:** Predicted volume obtained resampling the reference volume using the volume-wise motion estimates. **D:** Innovation image resulting from subtracting B from A. **E:** Innovation image resulting from subtracting C from A.

FIGURE 4.

Plot of the normalized mean squared error (NMSE) around a transition motion event (frame 270) showing how volume-wise motion correction cannot correct for the fast displacement until the movement has ended. A–F in the plot indicate several values of NMSE that are illustrated in the panels below. **A:** Original moving image; predicted volume using high-temporal motion estimates and subtracted image between the two before the transition motion (head stable at tilted position). **B:** Original moving image; predicted volume using volume-wise motion estimates and subtracted image between the two before the transition motion (head stable at tilted position). **C:** Original moving image; predicted volume using high-temporal motion estimates and subtracted image between the two during the transition motion. **D:** Original moving image; predicted volume using volume-wise motion estimates and subtracted image between the two during the transition motion. **E:** Original moving image; predicted volume using high-temporal motion estimates and subtracted image between the two after the transition motion (head stable at centered position). **F:** Original

moving image; predicted volume using volume-wise motion estimates and subtracted image between the two after the transition motion (head stable at centered position).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

TABLE 1

Mean and [standard deviation] of the NMSE metric between the original moving image and the predicted volume obtained applying volume-wise motion estimates (second column); and between the original moving image and the predicted volume obtained applying high-temporal motion estimates (third column).

Volunteer 1	Volume-wise estimates	High-Temporal estimates
Motion-free run 1	0.015 [0.003]	0.009 [0.003]
Staged continuous motion run 1 (z-rotation)	0.464 [0.199]	0.143 [0.040]
Motion-free run 2	0.095 [0.001]	0.088 [0.001]
Staged continuous motion run 2 (z-rotation)	0.384 [0.144]	0.130 [0.020]
Motion-free run 3	0.102 [0.002]	0.091 [0.001]
Staged continuous motion run 3 (z-rotation)	0.397 [0.144]	0.150 [0.021]
Volunteer 2	Volume-wise estimates	High-Temporal estimates
Motion-free run 1	0.006 [0.002]	0.011 [0.004]
Staged continuous motion run 1 (z-rotation)	0.179 [0.076]	0.060 [0.011]
Motion-free run 2	0.051 [0.007]	0.062 [0.004]
Staged continuous motion run 2 (z-rotation)	0.136 [0.045]	0.065 [0.004]
Motion-free run 3	0.045 [0.003]	0.061 [0.005]
Staged continuous motion run 3 (z-rotation)	0.141 [0.048]	0.068 [0.004]
Volunteer 3	Volume-wise estimates	High-Temporal estimates
Motion-free run 1	0.017 [0.004]	0.015 [0.004]
Staged continuous motion run 1 (x-rotation)	0.080 [0.019]	0.063 [0.015]
Motion-free run 2	0.074 [0.004]	0.072 [0.005]
Staged continuous motion run 2 (x-rotation)	0.096 [0.019]	0.080 [0.016]