

A quantitative comparison of motion detection algorithms in fMRI

Babak A. Ardekani*, Alvin H. Bachman¹, Joseph A. Helpert

Center for Advanced Brain Imaging, Nathan Kline Institute, 140 Old Orangeburg Road, Orangeburg, NY 10962, USA

Received 11 April 2001; accepted 17 June 2001

Abstract

An important step in the analysis of fMRI time-series data is to detect, and as much as possible, correct for subject motion during the course of the scanning session. Several public domain algorithms are currently available for motion detection in fMRI. This paper compares the performance of four commonly used programs: AIR 3.08, SPM99, AFNI98, and the pyramid method of Thévenaz, Ruttimann, and Unser (TRU). The comparison is based on the performance of the algorithms in correcting a range of simulated known motions in the presence of various degrees of noise. SPM99 provided the most accurate motion detection amongst the algorithms studied. AFNI98 provided only slightly less accurate results than SPM99, however, it was several times faster than the other programs. This algorithm represents a good compromise between speed and accuracy. AFNI98 was also the most robust program in presence of noise. It yielded reasonable results for very low signal to noise levels. For small initial misalignments, TRU's performance was similar to SPM99 and AFNI98. However, its accuracy diminished rapidly for larger misalignments. AIR was found to be the least accurate program studied. © 2001 Elsevier Science Inc. All rights reserved.

Keywords: Motion detection; Functional magnetic resonance imaging (fMRI)

1. Introduction

Functional magnetic resonance imaging (fMRI) is an *in vivo* imaging technique for the study of brain function. Very fast imaging pulse sequences, such as the echo-planar imaging (EPI), are used to scan a time-series of volumes (the term “volume” here refers to a set of 2D parallel image slices) from the brain. By studying the changes in the MR signal as a function of time in conjunction with sensory, motor, or cognitive stimulation, it is possible to localize regions in the brain that are “activated” as a consequence of the task performed by the subject [1].

The changes in the MR signal intensity that result from the subject stimulation are typically small as compared to the baseline signal intensity and noise level. Hence, the data can be easily corrupted by small involuntary movements by the subject between volume acquisitions. Therefore, it is essential to “correct” for the motion between volumes over the course of the fMRI experiment. The ideal approach to deal with motion is to detect it online and adjust the scan-

ning field of view before the next volume is acquired. While several such methods have been published in recent years [2–4], the technology is still not widely available. By far the most often used approach in dealing with the problem of motion in fMRI is retrospective image registration [5–8]. The problem of image registration can be separated into two parts: motion detection, and correction. Motion detection is the problem of estimating the set of translational and rotational parameters that give the exact position of the brain in 3-dimensional (3D) space for each volume in the time-series relative to its position at the start of the experiment or another reference time point. In the correction step, the estimated parameters are used to realign the volumes by interpolation. The focus of this paper is on motion detection. An excellent survey of the latter problem (interpolation) is provided by Lehmann et al. [9].

Several public domain programs are available that can be used for motion detection. These algorithms, when applied to common data sets, do not always produce the same results. This immediately raises two questions: (a) which algorithm gives more accurate motion estimation?; and (b) to what extent do the differences in the results affect the final activation maps? The objective of this study was to answer the first question, that is to evaluate the relative accuracy of these algorithms and to characterize their per-

* Corresponding author. Tel.: +1-845-398-5471; fax: +1-845-398-5472.

E-mail address: ardekani@nki.rfmh.org (B.A. Ardekani).

¹ Professor Emeritus, The City College/CUNY.

formance. Four programs that are commonly used by fMRI researchers for image registration were compared: SPM99 [5], AIR 3.08 [6], AFNI98 [8,10], and the pyramid approach by Thévenaz et al. [7] (TRU). A comparison of the first three algorithms has been previously published by Morgan et al. [11] based on the number of false positives each algorithm generated when used in analyzing simulated fMRI time-series with simulated activation foci. In this study, we compare the algorithms based on their actual accuracy in detecting the motion in simulated fMRI time-series with a range of known motions and different levels of added noise. A brief overview of the algorithms is presented in the following section. The subsequent sections include the methods used for generating the simulated data, results of applying the registration programs to the data, and conclusions.

2. Motion detection in fMRI

While many different and general image registration methods have been published, the algorithms for fMRI motion detection are usually designed to take advantage of three characteristics of fMRI time-series data: (a) the motion can be modeled as a rigid body motion in 3D described by a set of six parameters (three rotations about and three translations along the axes of a Cartesian coordinate system); (b) the image contrast does not change appreciably from one volume to the next; and (c) the motion is generally small compared to image resolution.

Let $f_n(\mathbf{r})$ represent the n th volume from a total of N volumes acquired during the fMRI experiment ($n = 1, 2, \dots, N$), where $\mathbf{r} = [x \ y \ z]^T$ is a position vector variable pointing to the space coordinates (x, y, z). Properties (a) and (b), above, imply the following model relating the n th volume to the first:

$$f_n(\mathbf{r}) = \alpha_n f_1(\mathbf{R}_n \mathbf{r} + \mathbf{t}_n) + e_n(\mathbf{r}) \quad (1)$$

where \mathbf{R}_n is a 3×3 orthonormal rotation matrix with a determinant of 1 (i.e., no reflection) fully characterized by three rotational parameters, \mathbf{t}_n is a 3×1 translation vector comprised of three translational parameters, and $e(\mathbf{r})$ represents noise. The factor α_n accounts for a possible global difference in intensity between the two volumes. The matrix \mathbf{R}_n and vector \mathbf{t}_n represent the motion from time point 1 to n . The objective of motion detection algorithms in fMRI is to estimate, as accurately as possible, the six parameters that specify \mathbf{R}_n and \mathbf{t}_n .

The strategy in almost all motion detection methods is to define a “cost function” $J(\theta_n)$ as a function of the unknown parameters, θ_n , and minimize this cost function with respect to θ_n . Discrepancies between the results of applying different methods to the same data can arise from two main sources: differences in their optimization strategies for finding the minimum of the cost function; and differences in the definitions of the cost functions and their sensitivities to

motion and noise. The cost functions of the four algorithms that are considered in this paper are all special cases of the least-squares cost function:

$$J(\theta_n) = \sum_{\mathbf{r} \in \Omega} \omega(\mathbf{r}) [f_n(\mathbf{r}) - \alpha_n f_1(\mathbf{R}_n \mathbf{r} + \mathbf{t}_n)]^2 \quad (2)$$

where θ_n represents the six rigid body motion parameters in \mathbf{R}_n and \mathbf{t}_n in addition to the scale parameter α_n , and Ω represents the regular grid of voxel positions. The function $\omega(\mathbf{r})$ is a weighting factor. Since all algorithms considered in this paper employ variants of the least-squares cost function, any discrepancies observed between their results are more likely due to differences in their optimization strategies.

3. Methods

3.1. Data acquisition

MRI scans were obtained using a Siemens 1.5 T Magnetom Vision system (Siemens AG, Erlangen, Germany) equipped with a quadrature transmit/receive RF head coil. A high resolution multislice transverse brain turbo spin echo (TSE) scan was acquired from a normal healthy volunteer. The subject was positioned supine on the scanner bed. Subject motion was restrained by placing foam pads around his head. The data consisted of 50 interleaved slices of 2 mm thickness with no gaps, approximately covering the entire brain. The image size was 256×256 pixels with a resolution of $1 \times 1 \text{ mm}^2$. The acquisition parameters were: TE = 96 ms, TR = 9.5 sec, $\alpha = 90^\circ$, and ETL = 7. The imaging time was approximately 4 minutes and 40 seconds. This scan was used to simulate an fMRI time-series with motion. This pulse sequence was selected in order to obtain an image contrast similar to T_2^* -weighted EPI images (Fig. 1a).

4. Motion simulation

In order to simulate an fMRI time-series with a wide range of motions, 3000 θ 's (sets of six rigid body motion parameters) were generated randomly by sampling from a uniform distribution with limits of $\pm 7.2^\circ$ on angles and ± 6.0 mm on translations. The TSE volume was thresholded at 20% of the maximum voxel intensity to isolate the brain. A test set of voxels consisting of approximately 1/16 of the brain voxels was selected. The test voxels on each slice were located on a grid of points at 16 mm intervals. Let $\Omega_0 = \{\mathbf{r}_i\}$ represent this test set. The average *misalignment* of voxels in Ω_0 was computed for each of the 3000 θ 's as follows:

$$d = \frac{1}{\text{card}(\Omega_0)} \sum_{\mathbf{r}_i \in \Omega_0} \|\mathbf{r}_i - \mathbf{R}\mathbf{r}_i - \mathbf{t}\|_2 \quad (3)$$

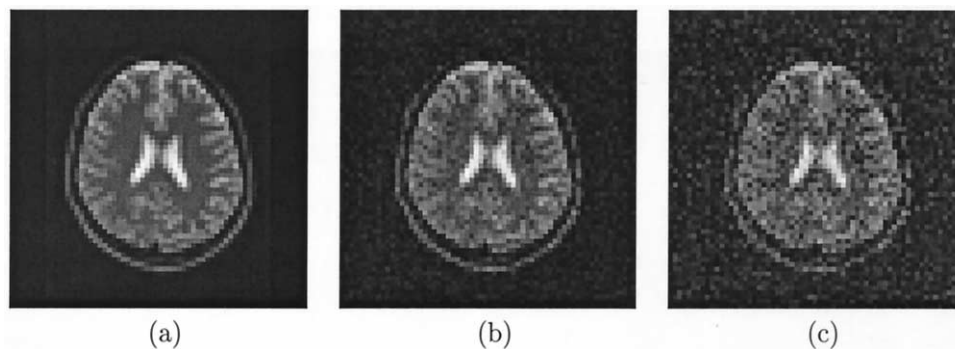


Fig. 1. A representative slice from the TSE volume after reduction to a $64 \times 64 \times 25$ resolution with (a) no added noise; (b) SNR = 10; and (c) SNR = 5.

where $\text{card}(\Omega_0)$ denotes the number of voxels in Ω_0 . The 100 parameter sets that produced misalignments closest to the 0.1 mm marks between 0.0 and 10.0 mm were taken as our test parameters. For each of the 100 θ 's thus selected, the original high resolution TSE volume was transformed and interpolated using the “3drotate” module of AFNI98. They were then down-sampled to a $64 \times 64 \times 25$ resolution with isotropic voxels of size $4 \times 4 \times 4 \text{ mm}^3$. Interpolation of the finer resolution TSE image followed by collection into coarser voxels reduces the interpolation errors in the simulated fMRI volumes. Thus, we obtained 100 volumes with progressively increasing misalignments from 0.1 mm to 10.0 mm. The original TSE volume was also down-sampled to the lower resolution, resulting in a total 101 volumes, where the exact motions of all the volumes with respect to the first volume were known. This data set was used to assess the accuracy of the motion detection algorithms.

4.1. Noise simulation

The 101 test volumes generated using the procedure outlined above were modified by the addition of Rician noise [12]. Four different signal-to-noise ratios (SNR) were simulated: SNR = 2, 5, 10, 20. The SNR was defined as A/σ , where A is the mean voxel intensity over the thresholded brain region and σ is the standard deviation of the noise in the real and imaginary channels. Since A is known, σ can be computed for a given SNR. For a given σ , the value of each voxel in the volume was replaced by a sample from the Rician probability distribution function:

$$p(v) = \frac{v}{\sigma^2} \exp\left[-\frac{v^2 + v_0^2}{2\sigma^2}\right] I_0\left(\frac{v \cdot v_0}{\sigma^2}\right) \quad (4)$$

where v_0 represents the initial voxel intensity in the absence of noise, v represents the random variable whose realization replaces v_0 , and I_0 is the modified zeroth order Bessel function of the first kind. A representative slice without noise, and at two noise levels (SNR = 5, 10) is shown in Fig. 1.

4.2. Experiments

The simulated fMRI time-series data were used to assess the accuracy of the motion detection algorithms. For each volume in the time-series, the *realignment error*, denoted by ϵ , was defined as the average displacement of the voxels in Ω_0 after image realignment. It is computed as follows:

$$\epsilon = \frac{1}{\text{card}(\Omega_0)} \sum_{\mathbf{r}_i \in \Omega_0} \|\mathbf{r}_i - \hat{\mathbf{R}}^{-1}(\mathbf{R}\mathbf{r}_i - \mathbf{t} - \hat{\mathbf{t}})\|_2 \quad (5)$$

where \mathbf{R} and \mathbf{t} are the known rotation and translation matrices, respectively; and $\hat{\mathbf{R}}$ and $\hat{\mathbf{t}}$ are their estimates obtained from the motion detection algorithm. Note that if $\hat{\mathbf{R}} = \mathbf{R}$ and $\hat{\mathbf{t}} = \mathbf{t}$, then the realignment error becomes zero.

Each algorithm was applied to the 4 noisy data sets with SNR = 2, 5, 10, 20 and to the data set with no noise (SNR = ∞). For each case, the realignment error ϵ was computed for each of the 100 misaligned volumes. The results are presented in the next section.

5. Results

Figure 2 shows the realignment errors (ϵ) of the four algorithms as a function of the initial misalignment (d). Each point in this figure is the average of 10 points in a 1 mm interval. That is, the realignment error at $d = 1.0$ represents the average of the 10 realignment errors at $d = 0.1, 0.2, \dots, 1.0$; the realignment error at $d = 2.0$ represents the average of the 10 realignment errors at $d = 1.1, 1.2, \dots, 2.0$, etc. The results shown in this figure are obtained from applying the four motion detection algorithms to the simulated data without noise (SNR = ∞). This exercise essentially tested the robustness of the algorithms with respect to the initial misalignment.

It can be seen that AIR provided the least accurate results over much of the range. TRU performed well when the initial misalignment was less than approximately 3 mm. The realignment errors in TRU and AIR depended strongly on the initial misalignments. They increased rapidly with in-

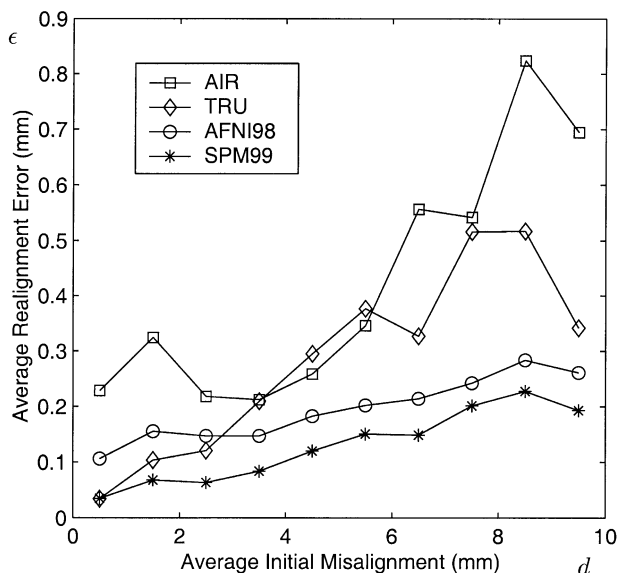


Fig. 2. The average realignment errors, ϵ , after applying the four motion detection algorithms to the simulated fMRI data with SNR = ∞ with initial misalignments, d , in the range of 0 to 10 mm.

creasing d . SPM99 provided the most accurate results followed by AFNI98. The accuracy of SPM99 and AFNI98 decreased slowly with increasing initial misalignment.

To evaluate the performance of the algorithms in presence of noise, we averaged the realignment errors in the range of 0–3 mm for noise levels SNR = 2, 5, 10, 20 and SNR = ∞ . In this range, for the noise-free case, all algorithms had performed relatively accurately. The results are shown in Table 1. As expected, the average realignment error increases as the SNR decreases. An exception was AIR where the average realignment errors for SNR = 10 was smaller than that of SNR = 20, which was in turn smaller than that of the noise-less case (SNR = ∞)! It can be seen that for large noise levels SNR = 2, 5, AFNI98 provided more accurate realignments than the other 3 programs, especially for SNR = 2 where there is a significant increase in the average realignment errors of SPM98, TRU, and AIR.

Table 2 shows the approximate processing times of the four programs on a SUN Ultra20 workstation. The processing times shown for AIR, AFNI98 and TRU include reslicing time. The processing time shown for SPM99 is for

Table 1
Average realignment error in mm over the range of initial misalignments from 0 to 3 mm for different noise levels

SNR:	∞	20	10	5	2
AIR	0.257	0.236	0.192	0.261	2.162
TRU	0.086	0.103	0.137	0.221	0.617
AFNI98	0.136	0.139	0.143	0.160	0.294
SPM99	0.055	0.062	0.091	0.216	0.885

The voxel size was 4 mm.

Table 2

Approximate processing times in minutes on a SUN Ultra20 Workstation for motion detection and reslicing of the 100 simulated misaligned volumes of size $64 \times 64 \times 20$

Program:	SPM99	TRU	AFNI98	AIR
Time (min.):	14*	20	4	19

*The time shown for SPM99 does not include reslicing.

motion correction only. AFNI98 was several times faster than the other three programs.

6. Conclusions

In this article, we applied the four motion detection programs: SPM99, AIR, AFNI98, and TRU to simulated fMRI data. Since the motion in the simulated fMRI data is known, we were able to assess the accuracy of motion detection in each of the algorithms. Based on the results obtained and presented in this article, several clear conclusions can be drawn. Amongst the four methods studied, AIR provided the least accurate results. SPM99 was found to provide the most accurate motion detection when SNR = 10, 20 or SNR = ∞ . For SNR = 2, 5, AFNI98 performed better than SPM99. TRU performed well when the initial misalignment was small but was clearly less accurate than SPM99 and AFNI98 for larger misalignments. The realignment errors of all algorithms increased with the initial misalignment. However, the increase was much more rapid in AIR and TRU. AFNI98 was several times faster than the other three algorithms. In conclusion, the results of this paper suggest that SPM99 should be the method of choice when it is known that the noise in the images is not unusually high and when speed is not critical. AFNI98 provides a good compromise between speed and accuracy. It also showed the most robustness with respect to noise.

Acknowledgments

The authors would like to thank Drs. Robert Bilder, David Guilfoyle, Jan Hrabe, Kevin Knuth, Kelvin Lim, and Michael Lipton for helpful discussions and Mr. Raj Sangoi for his help in data acquisition.

References

- [1] Cohen MS, Bookheimer SY. Localization of brain function using magnetic resonance imaging. *Trends in Neurosciences* 1994;17:268–77.
- [2] Derbyshire JA, Wright GA, Henkelman RM, Hinks RS. Dynamic scan-plane tracking using MR position monitoring. *J Mag Reson Imaging* 1998;8:924–32.

- [3] Lee CC, Grimm RC, Manduca A, Felmlee JP, Ehman RL, Riederer SJ, Jack CR. A prospective approach to correct for inter-image head rotation in fMRI. *Mag Reson Med* 1998;39:234–43.
- [4] Thesen S, Heid O, Mueller E, Schad LR. Prospective acquisition correction for head motion with image-based tracking for real-time fMRI. *Mag Reson Med* 2000;44:457–65.
- [5] Friston KJ, Ashburner J, Frith CD, Poline JB, Heather JD, Frackowiak RSJ. Spatial registration and normalization of images. *Human Brain Mapping* 1995;2:165–89.
- [6] Woods RP, Grafton ST, Holmes CJ, Cherry SR, Mazziotta JC. Automated image registration: I. General methods and intrasubject, intramodality validation. *J Comp Assis Tomog* 1998;22:139–52.
- [7] Thévenaz P, Ruttimann UE, Unser M. A pyramid approach to sub-pixel registration based on intensity. *IEEE Trans Image Proc* 1998; 7:27–41.
- [8] Cox RW, Jesmanowicz A. Real-time image registration for functional MRI. *Mag Reson Med* 1999;42:1014–8.
- [9] Lehmann TM, Gonner C, Spitzer K. Survey: Interpolation methods in medical image processing. *IEEE Trans Med Imaging* 1999;18:1049–75.
- [10] Cox RW. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Comp Biomed Res* 1996;29:162–73.
- [11] Morgan VL, Pickens DR, Hartmann SL, Price RR. Comparison of functional MRI image realignment tools using a computer generated phantom. In: *Proceedings of the 8th Scientific Meeting of the International Society of Magnetic Resonance in Medicine, Vol. 1. Denver, Colorado: ISMRM; 2000:848.*
- [12] Gudbjartsson H, Patz S. The Rician distribution of noisy MRI data. *Mag Reson Med* 1995;34:910–4.