PARTICLE FILTERING FOR SLICE-TO-VOLUME MOTION CORRECTION IN EPI BASED FUNCTIONAL MRI

Yu-Hui Chen, Roni Mittelman, Boklye Kim, Charles Meyer, Alfred Hero

University of Michigan, Ann Arbor, MI USA

ABSTRACT

Head movement during scanning introduces artificial signal changes and impedes activation detection in fMRI studies. The head motion in fMRI acquired using slice-based Echo Planar Imaging (EPI) sequence can be estimated and compensated by aligning the images onto a reference volume through image registration. Registering EPI images volume by volume fails to consider head motion between slices, leading to biased head motion estimates. Slice-to-volume registration is used to estimate motion parameters for each slice by more accurately representing the image acquisition sequence. However, it is prone to image noise and geometric distortion, resulting in high variance estimates. In this work, we propose a Gaussian particle filter based head motion tracking algorithm to reduce the image misregistration errors. The algorithm models head motion by using a dynamic state space model (SSM) that tracks and estimates the head motion for each slice. The head motion parameters are modeled by a random walk, and the Gaussian particle filter algorithm based on a dynamic state space model (SSM) that tracks and estimates motion parameters for each slice by more accurately following the EPI acquisition sequence. 

To deal with the above problem, the head motion should first be estimated and then used to correctly place fMRI image slices into the fMRI volume. Image registration is a favorable approach for head motion estimation since it does not depend upon complicated system settings or additional equipment, like video cameras or head markers [1–4]. We model the head motion by rigid body transformation and the transformation parameters are estimated by optimizing pre-defined image similarity measures, e.g., cross-correlation or mutual information [5], between functional and reference images. In [6], the head motion is estimated for each functional volume by registering the volumes to a reference volume. However, since the EPI images are taken slice by slice, stacking the slices directly and treating them as volumes neglects the head motion between consecutive slices within the same volume, i.e., inter-slice motion, as illustrated in Fig. 1.

Mapping-slice-to-volume (MSV) [7] proposed by Kim et al. is the first work to address the slice-to-volume registration approach. As compared to a volume-to-volume registration approach, slice-to-volume approach is capable of estimating and correcting the head motion for each slice by more accurately following the EPI acquisition sequence slice by slice. However, the main issue of this approach is computational: the image similarity measure may not be convex over the parameter space, and may have many local maxima in the presence of noise and inadequate image features. As usual, choosing the initialization for the optimization process’s convergence is essential for accurate registration.

In this work, we propose a head motion tracking (HMT) algorithm based on a dynamic state space model (SSM) that tracks and estimates the head motion for each slice. The head motion parameters are modeled by a random walk, and the Gaussian particle filter [9] is used to estimate the head motion according to the observed EPI slices. The main advantage of this approach is that it utilizes the

Index Terms— Multimodal image registration, mutual information, particle filter tracking, 3D brain motion tracking

1. INTRODUCTION

Brain activation studies are intended to identify specific regions in the brain that are involved with particular tasks. Detection of functional regions is most commonly performed by acquiring functional magnetic resonance imaging (fMRI) data using echo planar imaging (EPI) where the signal contrast is caused by the change of oxygenation in blood flow associated with local upstream neural activity. Typically in order to detect brain activation in this noisy environment, averaging of responses over several identical stimuli must be performed. The MRI scanner generates a series of images of the subject’s head section and the images are concatenated to form the 3D volume of each scan. The time-series of volumes captures the brain activity signal during the experiment, which serves as the starting point for brain activity analysis.

Ideally, each voxel in the volume time series records the signal evolving in time for a specific position. However, if the head of the subject moves during the scanning process, the time variation of voxel locations results in blurring or loss of signal and severe degradation of the fMRI image. This effect accumulates additional noise in the activation signal, impairing activity analysis accuracy. In experiments that require verbalized activation studies, the head cannot be fixed because the subject is required to speak during scanning. Therefore, some degree of head motion is inevitable even with cooperative subjects.

This research was partially supported by ARO MURI grant W911NF-15-1-0479 and NIH grant 2P01CA087634-06A2.
information from previous acquired slices to provide a good start-
ing point and effectively reduces the parameter search space in the
optimization process, resulting in much better registration accuracy.
The experimental results in Section 4 show that our approach outper-
forms other methods in terms of head motion parameter estimation,
and misregistration error for both synthetic and noisy real data.

The paper is organized as follows. In Section 2, we review back-
ground of the general image registration problem as well as the ex-
isting head motion correction methods. In Section 3, we describe our
Head Motion Tracking (HMT) algorithm and how it is used to
estimate the motion parameters. Section 4 shows the experimental
results of different approaches for synthetic and real data. Section 5
concludes this paper.

2. HEAD MOTION ESTIMATION BY IMAGE
REGISTRATION

The aim of image registration is to find a one-to-one transformation
Tθ that maps a reference image IR onto a target image IT, which
may come from different imaging modalities. This is done by opti-


\[ \hat{\theta} = \arg \max_{\theta} \mathcal{M}(I_T, T_\theta(I_R)), \] (1)

where \( T_\theta(.) \) is the transformation function which is parameter-
ized by \( \theta \). The parameterization of \( T_\theta \) could account for rigid
body displacement, local deformations, or other relative differ-
ences between the reference and target image volumes. For head
motion, a rigid body displacement parameterization is adequate:

\[ \theta = [\alpha, \beta, \gamma, dx, dy, dz], \]

where \( \alpha, \beta, \gamma \) are spherical Euler angles, and \( dx, dy, dz \)
are spatial positions defining the origin of the spherical
coordinate system. The image similarity measure is the mutual
information (MI), which has been widely applied to multi-modal
biomedical image registration [10].

The image acquisition process starts by collecting an anatomical
volume \( V_{anat} \) of the subject’s head using \( T_1 \)-weighted MRI [11],
which serves as the reference \( I_R \) for a functional MR image.
The functional MR images are acquired via multislice single-shot echo-
planar imaging (EPI) sequences acquired by \( T_2^* \)-weighted MRI,
which has significantly lower spatial resolution than \( T_1 \)-weighted
MRI. Let \( V = \{ V_m \}_{m=1}^M \) denote the set of collected EPI volumes,
where \( M \) is the total number of volumes acquired during the brain
scan session. Each of the EPI volumes is composed of a set of
EPI slices \( V_m = \{ S_{mn} \}_{n=1}^N \), where \( N \) is the number of slices per
volume. The head motion is estimated by registering the set of EPI
images \( V \) onto the anatomical volume \( V_{anat} \). There are two main
approaches that are commonly used to perform this multi-modality
registration:

Volume-to-volume Registration: Friston et al. [6] proposed to
estimate the head motion for each volume by registering the EPI
images volume by volume via the optimization:

\[ \hat{\theta} = \arg \max_{\theta} \mathcal{M}(V_m, T_\theta(V_{anat})). \]

This advantage of this approach is that the 3D volume contains abundant image features. However, since the
EPI images are acquired slice by slice, this approach is not able
to track significant movement occurring between each EPI slice.
As EPI slices are commonly acquired in interleaved fashion, the
typical time elapsed between adjacent slices can be as large as 1
second [8, 12]. Therefore, inter-slice head motion can be significant.

Slice-to-volume Registration: This method maps each indi-
vidual slice into the anatomical reference volume space as pro-
posed in [7, 13]. The motion parameters are estimated for slices
instead of volumes via the optimization:

\[ \hat{\theta} = \arg \max_{\theta} \mathcal{M}(S_{mn}, T_\theta(V_{anat})), \]

where \( T_\theta(.) \) is the function that interpolates
the anatomical volume into 2D section with the motion parameter
\( \theta \). This approach is capable of estimating and recovering the inter-slice
head motion. However, because each 2D EPI slice \( S_{mn} \) carries
less information than the 3D volume \( V_m \), it is important to couple
together the registration of successive EPI slices. The coupling of
successive EPI slices in the registration process constitutes the main
contribution of this paper.

3. HEAD MOTION TRACKING

3.1. Coordinate Transformation

Our head motion tracking algorithm adopts the slice-to-volume ap-
proach to estimate the head motion for each EPI slice. We formulate
this problem as an optimization and use a Gaussian particle filter
to initialize and track the rigid body motion parameters across EPI
slices. Let \( S = \{ S_t \}_{t=1}^T \) denote the set of acquired EPI slices re-
arranged in order of acquisition time, where \( T = MN \) is the total
number of slices in the experiment. Given the acquired EPI slices \( S \)
and the anatomical volume \( V_{anat} \), the aim of the tracking algorithm
is to estimate the head motion parameters at each time \( \{ \theta_t \}_{t=1}^T \).
Since we model the head motion as a rigid body transformation,
the parameter \( \theta_t \) has six degrees of freedom and can be represented as a
3 × 3 rotation matrix \( R_t \) and a translation vector \( q_t \). Let \( x_t \), \( u_t \) de-
note the 3D-coordinates in the reference and observation coordinate
systems. The conversion between the two coordinate systems can be
described as:

\[ (x_t - c) = R_t ((R_t x_0 + q_0) - c) + q_t, \]

where \( R_t \), \( q_t \) are fixed transformations introduced by coordinate
mismatch between the two MRI scanners, e.g., due to initial head
position difference, and \( c \) is the head rotation center which approx-
imately corresponds to the location of the cervical vertebræ. Note
that \( R_t \), \( q_t \), \( c \) are constant over time and only need to be estimated
once in the whole experiment. The method to estimate these param-
eters is discussed in Section 3.3.

3.2. Head Motion Tracking Algorithm

We use a state space model (SSM) [14] to describe the head motion,
where \( \theta_t \) denotes the rigid body parameters at time \( t \). The state
equation is modeled using a Gaussian random walk with covariance
matrix \( \Sigma_d \):

\[ \theta_{t+1} = \theta_t + u_t, \quad u_t \sim \mathcal{N}(0, \Sigma_d) \] (3)

Note that our HMT algorithm can also be applied with more gen-
eral head motion model, e.g., a kinematic model [15]. The acquired
EPI slice, called the observation in the sequel, is related to the state
through the quasi-likelihood function:

\[ p(S_t|\theta_t) = \frac{1}{Z} \mathcal{L}(\mathcal{M}(S_t, T^{\theta_t}_t(V_{anat}))), \]

where \( \mathcal{L}(\cdot) \) can be chosen as any function such that it is positive and
monotonically increasing (i.e., \( \mathcal{L}(x) \geq 0, \forall -\infty < x < \infty, x > y \Rightarrow \mathcal{L}(x) > \mathcal{L}(y) \)) and \( Z \) is a normalization coefficient that turns
the objective function \( \mathcal{L}(\cdot) \) into a conditional probability, which is
denoted \( p(S_t|\theta_t) \) and is a quasi-likelihood function of \( \theta_t \). Here
\( S_t = \{ S_j \}_{j=t-h}^{t+h} \) denotes the stack of slices over a length \( 2h + 1 \) time interval centered at time \( t \). If \( h = 0 \), \( S_t \) is reduced to a single EPI slice \( S_t \). The parameter \( h \) controls the trade-off between parameter estimator bias and variance. In the analysis reported below, we have used \( h = 1 \), which was found to achieve a good trade-off between these two factors.

The Kalman Filter [16] is the optimal minimum mean squared error estimator for a linear SSM. The extended Kalman filter (EKF) [17] and the unscented Kalman filter (UKF) [18] use some form of approximation to deal with non-linear cases. These approaches require explicit state and observation equations, which are not readily available in the fMRI problem treated here. Therefore, in this work we adopt the sequential importance sampling approach (e.g., the particle filter [19]) to approximate the posterior distribution of the state from the proposed quasi-likelihood function (4).

Our Head Motion Tracking (HMT) algorithm is based on the Gaussian particle filter (GPF) [9] framework that uses a set of weighted samples, called particles, to approximate the state and observation distributions. Initially slice-to-volume registration is used to generate an initial head motion estimate \( \hat{\theta}_0 \). As in the GPF, for each slice at time \( t \), the algorithm has two stages: Measurement update and Time update. In the Measurement update stage, we use \( P \) particles \( \{ \theta_k \}_{k=1}^{P} \) drawn at the last time step to evaluate the particle weights using the quasi-likelihood function \( p(S_t|\theta_k) \) defined in (4). The quasi-likelihood function should have two properties: (1) It is monotonically increasing with the image similarity \( M(S_t,T_\hat{\theta}(V_{anat})) \); (2) The weighted particles are approximately distributed according to a multivariate Gaussian density. To satisfy the two properties, we propose to use a histogram equalization approach to evaluate the particle weights. The target density is the distribution of \( z = f(x) \) where \( x \) and \( f(.) \) are the 6-dimensional multivariate Gaussian random variable and density, respectively. Letting \( g_\theta(z) \) denote the density of \( z \), we can equalize the histogram to obtain the particle weights.

\[
g_\theta(z) = \pi^3 (-2 \log (2\pi^3 z)^2), \quad z \in (0, (2\pi)^{-3}). \quad (5)
\]

The particle weights are normalized to sum to 1 and used to calculate the weighted mean and covariance. The weighted mean is then used as the initialization to optimize:

\[
\hat{\theta}_t = \arg \max_\theta M\pi(S_t,T_\theta(V_{anat})). \quad (6)
\]

The transformation parameter \( \hat{\theta}_t \) that maximizes (6) is the estimated head motion at time \( t \). Since the weighted mean incorporates abundant information about the image similarity distribution in neighboring regions, the optimization process can largely benefit from the good initialization. In this paper, we use the Nelder-Mead [20] optimizer, which is a simplex method used to iteratively find the optimum of an objective function in a multi-dimensional space. The mutual information is calculated using histogram approach [10], where the reference volume \( V_{anat} \) is first tri-linearly interpolated to the same resolution as the target image \( S_t \). Note that the proposed histogram equalization approach is not restricted to any particular definition of image similarity. Therefore MI can be replaced by any other image similarity measure, e.g., Normalized MI [21], localized MI [22], or graph-based MI [23]...etc. After the motion parameter is estimated, we perform a standard re-sampling step to estimate the covariance matrix of the posterior distribution, which is then used to establish the prior distribution of the next slice in the Time Update stage using (3).

In real data, sometimes the acquired images are very noisy and difficult to register, especially the lower and upper apex of the head. To reduce the effect of these noisy slices, we screen the slices for adequate signal strength. Specifically, we reject all EPI slices that have fewer than 15% of the pixels above a certain threshold value. For these rejected slices, we skip the optimization step and estimate the motion parameters through interpolation of the estimates from neighboring slices. We use 2nd-order interpolation, which is accurate when the head motion has approximately constant angular and translational velocities.

### 3.3. System Parameters Setting

In the proposed Head Motion Tracking algorithm there are several parameters that need to be set: \( R_x, q_x, c, \Sigma_x \).

**Fixed Coordinate Transformation \( R_x, q_x \):** Since \( R_x, q_x \) are constant over the entire experiment, they can be estimated by first taking the average of all EPI volumes over time, and then registering the averaged EPI volume to the anatomical volume to obtain as estimate of these parameters.

**Head Rotation Center \( c \):** To estimate the head rotation center, we run the HMT algorithm on the first \( K \) EPI slices (we used \( K = 70 \) in our experiment) by assuming \( c = 0 \) as the origin. Let \( \{ \hat{\theta}_i \}_{i=1}^{K} \) denote the estimates of the motion parameters for these \( K \) image slices. Here we assume that the patient’s body position is stable during the scan (the subject is tied and lying in the machine) and therefore the amount of translation should be small, i.e., \( \| q_i \| \approx 0 \). Based on this assumption, the rotation center should be the coordinate that minimizes the average amount of translation which can be estimated by solving the least squares problem:

\[
\hat{c} = \arg \min_c \sum_{i=1}^{K} \| q_i - (I_3 - R_t)c \|^2, \quad (7)
\]

where \( I_3 \) is the \( 3 \times 3 \) identity matrix.

**Head Motion Covariance \( \Sigma_x \):** The estimate of the head motion covariance matrix is generated in two steps. First, we initially set \( \Sigma_x \) to the identity and run the HMT algorithm over \( K \) image slices to obtain the estimates \( \{ \hat{\theta}_i \}_{i=1}^{K} \). Subsequently, the matrix \( \Sigma_x \) is estimated as the covariance matrix of the consecutive parameter differences: \( Cov(\hat{\theta}_t - \hat{\theta}_{t-1}) \).

### 4. EXPERIMENTAL RESULTS

#### 4.1. Synthetic Data Generation

We downloaded high resolution \( T_1, T_2 \)-weighted MRI volumes from the International Consortium of Brain Mapping (ICBM) [24]. The high resolution \( T_1 \) MRI brain volume was used as the anatomical reference volume with voxel size 0.78 × 0.78 × 1.5mm³. The EPI slices were emulated by interpolating the \( T_2 \)-weighted volume under artificial motion induced by setting the motion parameters as in [25]. The voxel size of the EPI slices is 1.56 × 1.56 × 6mm³, a blurring Gaussian low-pass kernel with \( \sigma = 2 \) was applied, and 3% Gaussian noise was added to simulate real EPI slices. This produced a synthetic EPI data set consisting of \( M = 120 \) volumes with \( N = 14 \) slices per volume.

We evaluate the performance quantitatively with respect to head motion parameter estimation error and misregistration error. The misregistration error is measured by average voxel distance, which is the average distance between the registered voxel coordinate and the true voxel coordinate. Let \( x^{reg}(i) \) and \( x^{true}(i) \) denote the coordinates of voxel \( i \) transformed using the estimated motion parameter
$\hat{\theta}_t$ and true motion parameter $\theta_t$ of slice $t$. The average voxel distance is defined as $D_t = 1/N_v \sum_{i=1}^{N_v} \|x^{\text{true}}_i - x^t_i\|$, where $N_v$ is the total number of voxels in a single slice.

4.2. Evaluation Using Synthetic Data

The simulated EPI slices described in Section 4.1 are registered to the anatomical volume to estimate the motion parameters by using the following three methods (implemented in MATLAB): (1) volume-to-volume registration [6] (V2V); (2) slice-to-volume registration [7] (S2V), where the optimization process is initialized by the V2V result; (3) the proposed Head Motion Tracking algorithm (HMT) with $P = 4000$ particles. Figures 2(a)-(c) show the motion parameters of three rotation angles for the first 200 slices, where the black lines denote ground truth and the dashed lines denote estimated motion parameters. Figure 2(a) demonstrates that the volume-to-volume registration method can accurately estimate motion for each volume but cannot accurately track the motion over the slices. On the other hand, S2V (Fig. 2(b)) can better track the head motion over different slices but has high bias, especially for slices near the apex of the head where slice image intensity and contrast are low. Our proposed HMT algorithm (Fig. 2(c)) is able to track the head motion much more accurately than the other two approaches. Figure 2(d) shows the boxplot of the average voxel distance after registration for different methods. The whiskers are the outliers outside the inner fence (defined by $1.5 \times F$-spread [26]). All of these methods reduced a fair amount of the voxel misregistration errors compared to no motion correction case (NoCorr). Notice that our HMT algorithm has significantly lower misregistration error, as measured by voxel distance. The mean of $D_t$ over all slices are listed above Fig. 2(d).

4.3. Evaluation Using Real Data

We further validate the performance of the proposed HMT algorithm on real human data. The real data was acquired from a normal volunteer. We asked the subject to intentionally nod his head during the EPI acquisition process. The head was scanned 126 times with 14 slices in each volume for this dataset. The anatomical voxel size is $1 \times 1 \times 1.5\text{mm}^3$ and the EPI voxel size is $2 \times 2 \times 6\text{mm}^3$.

Figure 3 shows the three Euler angles estimated by S2V (a) and HMT (b) overlaid with the V2V result (black lines) for the first 200 slices. Similarly to the experiments with synthetic data in Section 4.2, S2V can be used to estimate the motion for each slice but is noisy. The abrupt changes in the motion parameters demonstrated by S2V represent unlikely head movement, which suggests incorrect estimation. On the other hand, our HMT algorithm produced much more stable and smoother motion estimates, which is more convincing in describing real head motion. The superior performance of HMT is a consequence of the dynamical modeling that couples together estimates from successive slices leading to smoother and less noisy tracking performance.

5. CONCLUSION

In this work, we have proposed a head motion tracking (HMT) algorithm that uses an image registration objective function and the Gaussian particle filter to couple motion estimates from successive EPI slices, resulting in improved performance. Due to the fact that the proposed algorithm utilizes the information from consecutive observations to effectively improve the optimization process, it combines the bias reduction properties of the S2V approach and the variance reduction properties of the V2V approach. The experimental results demonstrated that the proposed MT algorithm can significantly improve the estimation accuracy over the volume-to-volume and slice-to-volume approaches in terms of motion parameter estimation and misregistration error in synthetic and real human experimental data.

6. REFERENCES


