Review of Preprocessing Steps for BOLD fMRI

Article

Evaluating fMRI Preprocessing Pipelines

Review of Preprocessing Steps for BOLD fMRI

BY STEPHEN C. STROTHER

During the past decade, many papers have been published that describe image processing of functional magnetic resonance image (fMRI) time series. These papers mainly focus on processing real-valued, magnitude images that are reconstructed from *k*-space data (i.e., fourier space), before application of a statistical analysis approach to detect brain activations. These processes have come to be known collectively as preprocessing operations (Figure 1). The sequential combination of spatio-temporal image processing steps following data acquisition and including data analysis is referred to as the fMRI data-processing pipeline [1]. There has also been a recent focus on using complex-valued fMRI time series to improve activation detection [2]–[4] and remove vascular artifacts [5]. Most of the preprocessing issues associated with the use of complex-valued data have yet to be studied, and this approach is not discussed in the following. In addition to fMRI images, structural magnetic resonance images (MRIs) are passed to the preprocessing pipeline. These provide an anatomical reference for the statistical parametric images (SPIs) output from the data analysis stage (Figure 1) and allow multiple subjects to be registered within a common coordinate system. The existence of a distinct set of preprocessing steps has recently been codified in a separate chapter of the excellent introductory book on fMRI by Huettel et. al. [6]. Previous books covering fMRI and its applications include chapters that cover many of the preprocessing steps described in the following [7], [8].

In this article, I review the evaluation and optimization of the preprocessing steps for blood-oxygenation-level-dependent (BOLD) fMRI. This technique indirectly measures changes in local neuronal firing rates by measuring associated changes in deoxy-hemoglobin concentrations in nearby blood vessels. I will not directly address preprocessing operations for other MRI techniques (such as arterial spin labeling for measuring blood flow in the brain), structural MRI techniques (such as diffusion tensor imaging for measuring white matter tracts), or voxel-based morphometry for general analysis of structural MRI images.

By virtue of their central position in the processing pipeline, the preprocessing steps interact with virtually every decision made in designing and performing an fMRI experiment (Figure 1). For example, at the start of the pipeline, the theories of brain function and disease that are to be tested will determine the experimental design variables, which will in turn guide the choice of scanner pulse sequences, within the physical constraints of the scanner. This has led to several classes of fMRI experiments that are defined relative to the time during which the BOLD fMRI signal evolves, i.e., the hemodynamic response function (HRF). Following a single brief stimulus, the HRF lasts approximately 10–15 s. Block designs continuously present a stimulus for a single experimental condition over an extended period of time relative to the HRF (typically 20–30 s). Event-related designs present discrete, stimulus events with durations and interstimulus intervals that are typically short compared to the HRF and may be randomized; mixed designs contain features of both block and event-related designs. Many of these experimental design choices interact with, and therefore should influence, the choice of preprocessing steps and parameters.

In addition, there are two distinctly different standard pulse sequences for acquiring each slice of *k*-space data, echo planar imaging (EPI) that fills *k*-space on a rectangular grid, and spiral imaging (SI) that fills *k*-space along a corkscrew trajectory. These sequences produce somewhat different imaging artifacts and signal-to-noise properties, but differences in their interactions with preprocessing parameter choices have not been systematically studied. At the other end of the processing pipeline, the choice of data analysis approach may strongly interact with the preprocessing parameter choices, and these interactions are gradually being studied [1], [9]–[11].

The magnetic field strength also plays an important role in processing pipeline decisions. Many groups are upgrading from 1.5 T to 3.0 T, and even 7.0 T fields or higher for research scanners. It is well established that 3.0 T fields have a significantly higher contrast-to-noise for the BOLD effect than 1.5 T fields in homogeneous brain tissue [12]. However, this advantage is somewhat reduced by a higher physiological noise fraction, greater artifacts at air-tissue boundaries, and reduced decay times [13], i.e., preprocessing becomes more important with increasing field strength.

Because of these and many other interactions, the steps in the preprocessing pipeline should not be studied in isolation. Moreover, much of the preprocessing research outlined in the following has focused on developments driven

 0739-5175/06/$20.00©2006IEEE

The preprocessing steps interact with virtually every decision made in designing and performing an fMRI experiment.

by technological and methodological concerns (Figure 1) in which new approaches have been introduced for one or two of the preprocessing pipeline steps. These reports are typically accompanied by limited test data that demonstrate that the new approach is better than one or more of the existing approaches for a particular step. The reports often ignore possible interactions with other preprocessing pipeline steps or the task for which the pipeline was assembled in the first place. Such methodological approaches are important to understand the properties of algorithms and software for individual steps but do not guarantee the improvement of the generalizable and reproducible models of brain function and disease required at the output of the fMRI pipeline (Figure 1).

In all but a small number of large research centers, most fMRI data are collected with a predefined EPI or SI acquisition sequence and are often analyzed within a software package based on a predefined set of parameters that specify the preprocessing steps. Many fMRI studies in the literature are based on an analysis using the general linear model (GLM) within the statistical parametric mapping (SPM) software package [14] (Table 1). This observation is supported by the fact that 67 of the 100 distributable, published datasets available from the fMRI Data Center [15] were originally analyzed using some version of SPM [16]. As a result, it is preprocessing parameters within SPM or one of the other fMRI processing packages that are readily accessible to researchers and define the majority of processing pipelines that are used in the field. In this article SPM96, SPM99, and SPM2 refer to versions of the SPM software package. One of the challenges facing both academic researchers interested in fMRI processing pipelines and general users of the major software packages (Table 1) is the difficulty in identifying the current preprocessing pipeline practices in most major research groups. This information only partly exists in the academic literature and the e-mail discussion lists and user manuals of major software packages.

There is an ever-expanding collection of techniques and software tools available with which to assemble and apply many different data-preprocessing pipelines to fMRI datasets. One of the challenges facing the functional neuroimaging community is how to choose from among this combinatorial explosion of possible pipelines. This poses a range of critical issues, not the least of which is the day-to-day choice of which tools/packages to use and/or download and learn to use. All too often, this choice is resolved expediently in favor of availability, familiarity, and ease of use. I hope this review of the preprocessing literature will help to alert researchers to new approaches and possible preprocessing pipeline options that might be important to their research.

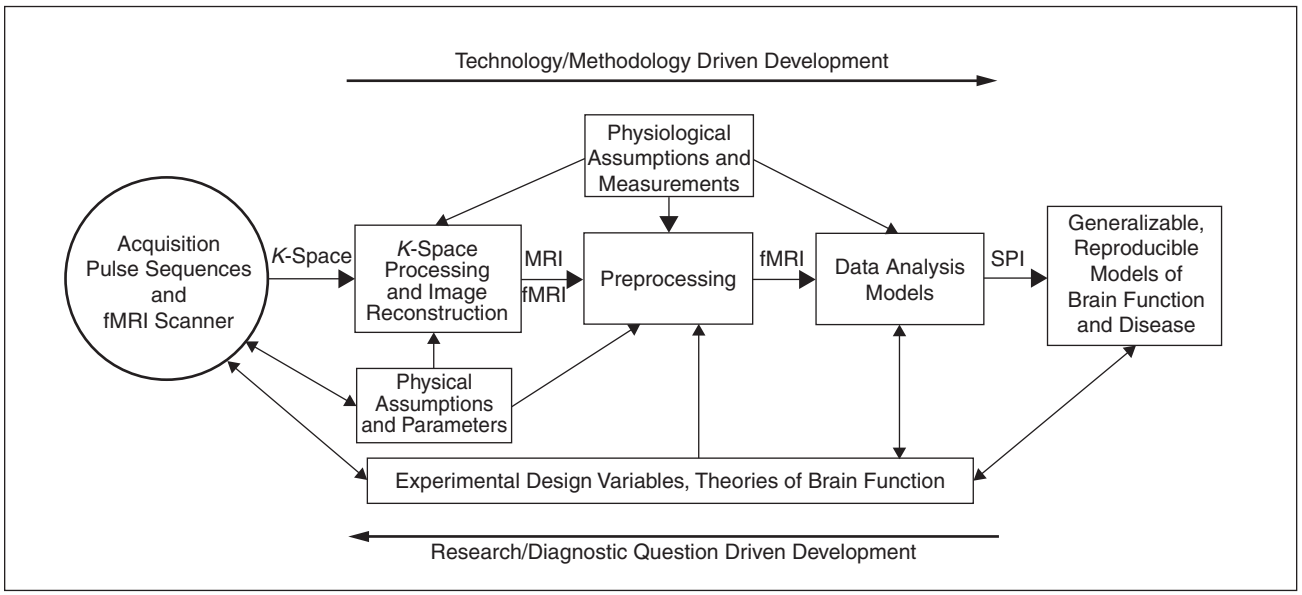

Fig. 1. A schematic representation of the flow of data through the fMRI processing pipeline from the acquisition of raw, *k*-space data, reconstruction by inverse Fourier transform to MRI and fMRI images, through the preprocessing pipeline and finally data analysis to produce statistical parametric images (SPIs) for identification of activated brain regions. Double-headed arrows indicate constraints and parameters moving in both directions.

There is an ever-expanding collection of techniques and software tools available with which to assemble and apply many different data-preprocessing pipelines to fMRI data sets.

The Steps of an fMRI Preprocessing Pipeline

Figure 2 illustrates the steps within the fMRI preprocessing pipeline defined by Figure 1. The individual steps are discussed in the following under their separate headings—except for quality control (QC), which is a sometimes neglected but an extremely important element of a well-executed preprocessing pipeline. Each of the QC stages QC2–QC5 will be discussed as part of the preceding processing step.

Quality Control 1

This first stage is perhaps the most critical, and it embodies the general principle that it is vital to look at the raw image data using viewing/survey tools, e.g., those provided in the major software packages. Performing QC1 is critical to avoiding running entire preprocessing and data analysis pipelines on poor-quality data. Unfortunately, because three-dimensional (3-D) fMRI image volumes, or scans, are typically collected one two-dimensional image slice at a time, they are susceptible to individual slice artifacts due to timing errors and radio frequency spikes. It is not uncommon on some scanners to see one or more slices in a single image volume that have some form of image artifact while the majority of the slices are of good quality. An example is shown in Figure 3, where a few adjacent volumes from several hundred had major artifacts in individual slices in the lower half of the brain.

Given the thousands of slices that may need to be viewed, the major software packages have all implemented cine viewers with which many slices may be rapidly reviewed using the eyes' sensitivity to dynamic changes to detect anomalous slices. This visual review should be supplemented with an exploration of the spatio-temporal structure of the multivariate time series of each subject's data. A number of simple approaches, such as plotting the time series of slice means, are discussed in [6], and many of these are available as supplemental software tools for the major packages, such as in the SPM2 diagnosis toolbox [17]. An alternative approach involves performing a principal component analysis (PCA) [18] or independent component analysis (ICA) [19] of the spatio-temporal time series from each subject. For example, both PCA and ICA can be run within the Multivariate Exploratory Linear Optimized Decomposition into Independent Components (MELODIC) tool in the Functional MRI of the

Table 1. General lists of fMRI software tools.

Bibliography on Neuroinformatics	http://www.imm.dtu.dk/~fn/bib/Nielsen2001BibNeuro informatics/node8
Analysis and Processing: Tools[1]	html#SECTION00081000000000000000
Internet Analysis Tools Registry[2]	http://www.cma.mgh.harvard.edu/iatr/display.php?spec=all
Some Software Tools with Flexible Preprocessing Pipelines and Data Analysis Models	
Analyses of Functional Images (AFNI)[3]	http://afni.nimh.nih.gov/afni
Brainvoyager (BV)[4]	http://www.brainvoyager.com/
FMRIB's Software Library (FSL)[5]	http://www.fmrib.ox.ac.uk/fsl/
Statistical Parametric Mapping (SPM)[6]	http://www.fil.ion.ucl.ac.uk/spm/
Selected Sites That Specify Default Processing Pipeline Choices	
SPM99 and SPM2 processing pipelines[7]	http://www.mrc-cbu.cam.ac.uk/Imaging/Common/fmridefaults.shtml
SPM99 fMRI processing pipelines[8]	http://www-psych.stanford.edu/~kalina/SPM99/Protocols/spm99_prepros_ prot.html
SPM99 fMRI processing pipelines[9]	http://psyphz.psych.wisc.edu/~oakes/spm/SPM99_demo_fmri.pdf
An AFNI processing pipeline[10]	http://brainimaging.waisman.wisc.edu/~tjohnstone/AFNI_I.htm
Neuroimaging Workflows in Fiswidgets[11]	http://grommit.lrdc.pitt.edu/fiswidgets/flow_doc/index.html

[1] *Maintained by Dr. Finn Nielsen in the Informatics and Mathematical Modeling group of the Technical University of Denmark.*
[2] *Maintained by the Centre for Morphometric Analysis at Massachusetts General Hospital, Harvard University.*
[3] *Free research software from the Scientific and Statistical Computing Core of the National Institute of Mental Health, National Institutes of Health, Bethesda, Maryland.*
[4] *Commercial software from Brain Innovation, B.V., Maastricht, The Netherlands.*
[5] *Free research software from the Oxford Centre for Functional MRI of the Brain (FMRIB), Dept. Clinical Neurology, John Radcliffe Hospital, Oxford University.*
[6] *Free research software from the Functional Imaging Laboratory, Welcome Department of Imaging Neuroscience, Institute of Neurology, University College London.*
[7] *MRC Cognition and Brain Sciences Unit (CBU) fMRI processing web pages maintained by Dr. Matthew Brett.*
[8] *Developed by Dr. Kalina Kristoff while a Ph.D. student and postdoctoral fellow in the Psychology Department, Stanford University.*
[9] *Developed by Dr. Terry Oakes, Director of Image Analysis and Visualization, The Waisman Laboratory for Brain Imaging and Behavior, University of Wisconsin—Madison.*
[10] *Developed by Dr. Tom Johnstone, The Waisman Laboratory for Brain Imaging and Behavior, University of Wisconsin—Madison.*
[11] *Processing pipelines that may be downloaded and run within the neuroimaging workflow management software, Fiswidgets (143).*

Brain (FMRIB) Software Library (FSL). The recent results from Beckmann and Smith [20] using MELODIC indicate that their probabilistic ICA (PICA) may have a number of advantages over standard ICA and PCA for both identifying and removing artifacts. Figure 4 illustrates PICA's ability to uncover many potential artifactual time series as separate components from a single fMRI dataset. One of the options with PICA is to filter the dataset by removing user-identified PICA components. This is likely to be a powerful technique for identifying and removing unwanted effects. A preliminary investigation of denoising using PCA and ICA has been performed by [21].

Another critical QC issue is image orientation, particularly absolute identification of the left and right hemispheres of the brain in an fMRI image volume. This problem is almost impossible to resolve post-hoc because of the left-right symmetry of normal brains. When setting up a new fMRI processing pipeline or changing any element in it, the pipeline should be retested with a phantom image containing absolute left-right labels. A quick search of the e-mail discussion lists of any of the major software packages reveals that this issue has been a constant problem. This problem occurs partly because there are two image display conventions in medical use: radiological or LAS (left brain = +x; anterior brain = +y; superior brain = +z), and neurological or RAS (right brain = +x; AS is the same as described for LAS). These conventions evolved because radiologists are taught to view images with left-right swapped as though viewing a supine subject from the feet (i.e., image left is subject right), and neurologists retain a subject's left-right orientation as though viewing from the head (i.e., image left is subject left). The most common fMRI file format, ANALYZE (http://www.mayo.edu/bir/PDF/ANALYZE75.pdf) assumes an LAS convention and does not formally support the neurological RAS convention. Unfortunately, the standard structural templates used by most groups (see "Registration III: MRI-Template") use the RAS convention, and this confusion has been dealt with differently by different software packages.

To deal with this problem, a new format for fMRI has been developed by the Data Format Working Group (DFWG), a committee set up under the Neuroimaging Informatics Technology Initiative (NIfTI, http://nifti.nimh.nih.gov/dfwg), which is sponsored by the National Institutes of Mental Health, Neurological Disorders and Stroke, and Biomedical Imaging and Bioengineering at the National Institutes of Health in the United States. The committee developed the NIfTI-1 format to provide an ANALYZE-compatible replacement for the current range of ANALYZE variants, both for file interchange and data storage. The four major fMRI data analysis packages in Table 1 are members of the DFWG and have agreed to use this format in the future with support for both input and output. NIfTI-1 resolves the LAS-RAS confusion by explicitly storing a voxel-index to spatial-coordinate transformation, typically based on scanner coordinates, in addition to a general affine transformation [22]. For an excellent summary of this state of affairs with additional useful URL links, see the SourceForge Web page contributed by Darren Weber (http://eeg.sourceforge.net/mri_orientation_ notes.html).

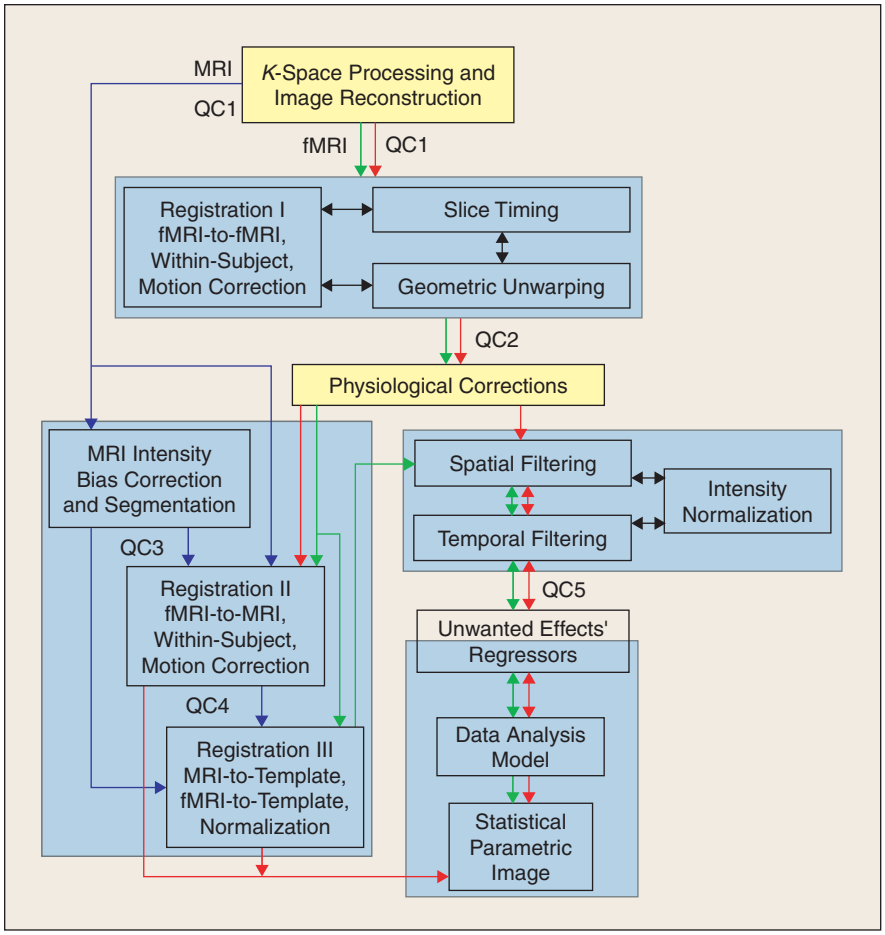

Fig. 2. A schematic representation of data flow through the preprocessing pipeline. Blue arrows represent structural MRI data, green arrows show the all-in-one approach, and red arrows illustrate the summary statistic approach to traversing the preprocessing steps that are discussed in the text. Gray boxes group steps that have conceptual similarities and/or strongly interact and may therefore be treated together, algorithmically and/or conceptually, as one combined processing stage in the pipeline. The QC points discussed in the text are marked between steps as QC1–5. Double-headed arrows depict strong interactions such that in some pipelines these steps will be combined into a single algorithm.

Registration I: fMRI-fMRI, Slice-Timing and Geometric Unwarping

Assuming the reconstructed images pass QC, the next step is to spatially align voxels across the sequentially collected fMRI image volumes. This ensures that each voxel's time series is an accurate representation of the BOLD time series for a voxel at a constant spatial location in the brain. The primary sources of time series

inaccuracies are 1) physical head motion, 2) within-volume slice-timing differences, 3) geometric distortions due to magnetic field inhomogeneities caused by susceptibility effects, and 4) residual-movement artifacts left after standard motion correction techniques and stimulus-coupled motion. Susceptibility artifacts are caused by the relatively large change in local magnetic field strength that occurs at a boundary between materials with very different magnetic susceptibility properties, e.g., an air- or bone-tissue boundary in the brain. Figure 4 illustrates several PICA components that relate to these artifacts. Figure 4(a) demonstrates typical edge artifacts seen as a consequence of uncorrected head motion. Figure 4(c) demonstrates signal fluctuations close to the sinuses, which are typical of the patterns resulting from the interaction of head motion and susceptibility-induced field inhomogeneities. Figure 4(d) illustrates phase artifacts, which can occur in the phase-encoding direction of slice acquisition [6].

Physical head motion cannot be completely eliminated while scanning, but it may be considerably reduced with careful use of head-immobilization techniques when placing the subject in the scanner and with training in an MRI simulator to accustom the subject to the scanning environment [6]. Using an MRI simulator, Seto et al. [23] have demonstrated that a subject's demographic group (i.e., stroke versus age-matched controls versus young adults) is more important than the type of motor task being performed in determining the range of head movements likely to be encountered in an fMRI study.

There is a large body of literature on retrospective correction of rigid-body head motion in fMRI, with each of the major packages having implemented their own approach: AFNI, [24]; FSL, motion correction by FMRIB's linear registration tool (MCFLIRT), [25]; SPM, [26], [27]; and BrainVoyager (the approach used seems to be unpublished). In addition, the automatic image registration (AIR) package developed by Roger Woods is widely used [28]–[30] (http://bishopw.loni.ucla.edu/AIR5/index.html). All of these approaches assume that the stack of slices comprising a volume are collected instantaneously, and that only rigid movement of the entire stack occurs across sequentially collected scans, i.e., a rigid-body assumption. Some of these algorithms have been compared by Ardekani et al. [31], who tested AIR3.08-linear, SPM99, and AFNI98; Morgan et al. [32], who tested AFNI96 and AFNI98, SPM96 and SPM99b, and AIR3.08 with linear and third-order polynomial registration; and Jenkinson et al. [25], who tested AIR3.08-linear, SPM99, and FSL's MCFLIRT.

These papers illustrate several of the difficulties in evaluating preprocessing pipelines based on the existing literature.

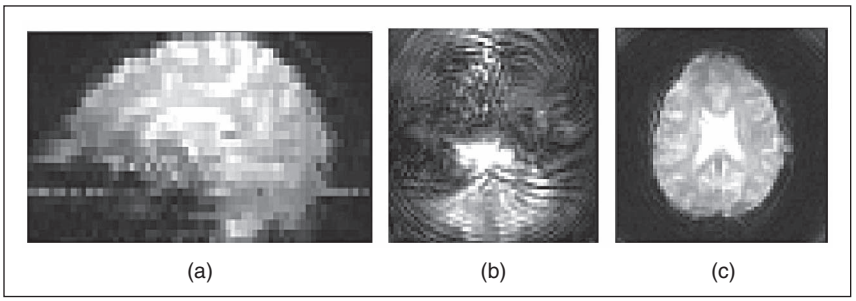

Fig. 3. Slice-dependent artifacts that can occur in a single volume image of an fMRI time series as a result of transient scanner instabilities during sequential slice collection. (b) A severely corrupted spiral acquisition slice from the (a) lower half of the brain (see saggital image) compared with (c) a normal slice from the upper half of the brain.

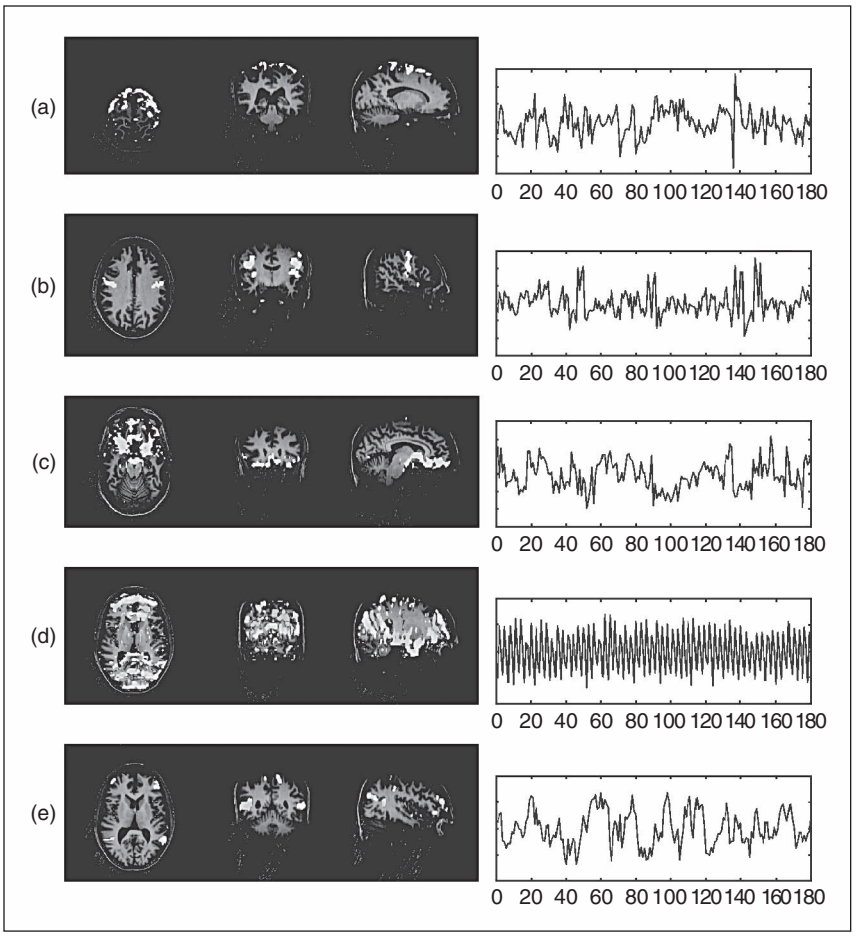

Fig. 4. Components from a probabilistic independent component analysis (PICA) depicting associated spatial maps (left) and time courses (right): (a) head motion (translation in Z); (b) sensory motor activation; (c) signal fluctuations in areas close to the sinuses (possibly due to interaction of B field inhomogeneity with head motion); (d) high-frequency MR ghost; and (e) resting-state fluctuations/physiological noise. The image is reproduced courtesy of Beckmann and Smith (20) from *IEEE Transactions on Image Processing* which contains further details.

First, there has never been a test of all five registration techniques from the four software packages in Table 1 and AIR, and these five do not constitute all of the registration algorithms in use and being developed (e.g., [33], [34]). Secondly, the software revision level is not always specified, and even if it is it may be difficult to establish the exact form of the previously published registration algorithm that is currently being used in a particular package. In addition, most packages provide many options, the settings of which should be, but are rarely, listed in the literature.

Important sources of error in retrospective algorithms designed to correct for physical head motion are the choice of the cost function for iterative algorithmic optimization and the interpolation technique used to resample the fMRI volumes and remove the measured movement effects.

Cost Functions

A cost function measures the similarity of each image volume in a time series to a reference volume, usually chosen as one of the scans in the time series, e.g., the first or the middle volume. The cost function is used to determine when an algorithm has generated sufficiently similar volumes to provide optimal estimates of the six rigid-body movement variables (three rotation and three translation). Many of the differences between the algorithms relate to the choice of a cost function and the optimization strategies implemented to iterate to a global, as opposed to a local, minimum. It has been clear for some time that intensity-based cost functions outperform landmark and fiducial-marker-based approaches [35], [36]. Much work during the last decade has indicated that cost functions other than a difference-of-squares perform best, although there is no consensus about any one optimal cost function. For example, Freire and Roche [37] found that mutual information and robust Geman-McClure estimators performed better than mean square error or correlation ratios in minimizing interaction with brain activation signals. Jenkinson et al. [25] found that correlation-ratio-based measures outperformed mutual information and least squares but did not consider interactions with brain activations.

Freire and colleagues have demonstrated that the standard retrospective alignment algorithms may even create spurious activations, particularly with least squares cost functions [38]. In addition, they have shown that more robust cost functions (i.e., the Geman-McClure M-estimator) reduce but do not eliminate the influence of task activation on estimated motion variables [37]. Two possible approaches to solving this problem using modified computational frameworks with least squares cost functions have been proposed and are compared in [34]. They find that both the down-weighting of the activated voxels that potentially bias motion variable estimates and simultaneously solving for motion and activation components in a GLM framework provide estimates that are robust to the effects of task activation.

Interpolation

Using interpolation techniques, new values are created (i.e., resampled) at new spatial locations. This is done using the six measured movement variables to define an interpolation function as a mathematical combination of the original voxel values in the spatial neighborhoods of the new locations. The type of interpolation scheme used represents a tradeoff between residual interpolation errors and speed. The ideal scheme for performing these interpolations on band-limited images, without introducing artifacts, is Fourier interpolation [24], [39]. In the spatial domain, this is equivalent to full-sinc interpolation, which uses every voxel in the image to calculate each new voxel value, a very slow and impracticable procedure for 3-D images. In order to gain speed, a truncated sinc is often used that needs values from only the local neighborhood of the new voxel location [40]. For AIR, Woods recommends a chirp-z interpolation within plane and a linear interpolation between planes [28]. Whether one of these schemes or some other is used, most interpolation approaches leave residual errors in the realigned volumes [28], [43] (trilinear interpolation leaves much larger residual errors than sinc-related and B-spline schemes [41], [42]). For example, Strother et al. [1] have demonstrated that using trilinear instead of windowed-sinc interpolation in a group-fMRI analysis is equivalent to additional in-plane smoothing with a Gaussian FWHM of approximately 1.5 pixels. Because interpolation techniques tend to introduce additional, spatially varying data smoothing, it is usually better to combine multiple resampling transformations (e.g., fMRI-fMRI and fMRI-MRI) and apply just one resampling and interpolation step.

Slice Timing

The second problem causing inaccuracies in time series is the sequential collection of slices within each volume, i.e., slice timing (Figure 2). While a new volume may be collected every 2–3 s (defined by the so-called acquisition TR), the individual slices are collected sequentially during this time period, sometimes in an interleaved manner, i.e., all odd slices are collected first followed by all even slices. For the usual ascending or descending sequential acquisitions, the last slice is collected almost one TR after the first slice; and with interleaved acquisitions, adjacent slices are collected a full TR/2 apart. Slice-timing correction uses interpolation between the same slice and voxel in neighboring acquisition TRs to estimate the signal that would have been obtained had the slices been acquired at the same time. The interpolated time point is typically chosen as the TR/2 time to minimize relative errors across each TR [6].

Any movement during the TR for one volume will affect the individual slices differently so that there is an interaction between movement, the rigid body assumption, slice-timing correction, and the type of slice acquisition sequence. A separate correction for slice-timing with temporal interpolation is provided by the major software packages, but there are no clear rules regarding whether this should be applied before or after rigid-body movement correction.

There appears to be a general consensus that slice-timing correction is probably unnecessary for most block designs, although this has not been clearly demonstrated in the literature. Huettel et al. [6] have proposed commonsense rules for single-event designs if the timing and registration corrections are performed separately: slice-timing correction should precede motion correction for an interleaved slice acquisition with a long TR; with sequential acquisitions or short TRs, motion correction should be done before timing correction.

There are two experimental approaches that integrate slice-timing corrections with slice registration: Kim et al. [44] introduced the mapping of single fMRI slices to a reference MRI collected in the same scanning session, the so-called map-slice-to-volume (MSV) approach, and Bannister et al. [45]

have introduced a preliminary model of constrained slice motion between sequential volumes. The MSV approach is discussed further under "Registration II:fMRI-MRI."

Susceptibility Artifacts

The third cause of motion-dependent inaccuracies in voxel time series is the interaction of head motion with magnetic field inhomogeneities, which are primarily caused by air- and bone-tissue magnetic susceptibility gradients, particularly at higher field strengths. A number of direct acquisition schemes exist to reduce this artifact, including static gradient shimming and modified pulse sequences, but appreciable image distortion typically remains in fMRI images [6]. In addition to fixed effects resulting from the static magnetic field, a voxel moving across an air-tissue boundary due to physical motion is further spatially displaced and distorted as a result of geometric effects due to its changing position in the local susceptibility gradient. This results in a dynamic distortion component as a result of movement. Such susceptibility gradients are most prominent in the frontal and temporal lobes with the boney, air-filled sinuses, causing serious artifacts in measurements of the orbito-frontal lobes of the brain [e.g., Figure 4(c)]. The effect may be somewhat reduced by acquiring smaller voxels to limit the gradient change per voxel, but this comes with tradeoffs of increased scan time and reduced signal to noise/voxel. Two retrospective, geometric unwarping approaches have been developed. The first and most common method used involves measuring the magnetic field inhomogeneities to create a field map for the fixed component, which is used to modify the acquired fMRI images [46], [47]. Attempts have also been made to estimate and use dynamic field maps collected at each time point [48]. If motion correction and geometric unwrapping are performed separately, Huettel [6] recommends performing rigid-body alignment before magnetic field distortion correction. This ordering is supported by Cusack et al. [49] and Elliot et al. [50], who have recently shown that it improves registration of motion-corrected, EPI fMRI images to structural MRI images in 3 T and 4 T scanners, respectively. The second retrospective method involves simultaneous, direct estimation of the rigid-body transformations for motion correction, together with a geometric unwarping correction, and this technique has been included in the SPM2 software package [51].

Residual-Movement Artifacts and Stimulus-Correlated Motion

The final sources of motion-dependent errors are residual-movement artifacts and stimulus correlated motion.

It is well-established that resampling using rigid-body transformation variables is insufficient to remove all motion effects from fMRI time series. Friston et al. [52] suggested that this is due to spin-history effects, and Woods et al. [28] noted that "motion-correlated artifact is commonly the greatest source of global variance in the motion-corrected dataset" and suggested that these effects are most likely due to residual interpolation effects (see also [43]). All of these investigators proposed removing these remaining movement effects by performing a GLM analysis using residual motion estimates as unwanted-effects regressors (Figure 2). A variety of choices are available for the regressors: the raw rigid-body movement variables, the first few components of their PCA [28], or sine and cosine basis functions with a one-voxel cycle [43]. Such residual-movement regression must be performed with care in the presence of stimulus-correlated motion as outlined in the following.

It is not uncommon, particularly in motor tasks, for subjects to produce small head movements that are correlated with performing the task. Bullmore et al. [53] provide a striking demonstration of the impact of stimulus-correlated motion when comparing schizophrenics, who are more susceptible to this confound than normal controls. They demonstrate that the standard motion-regression approach of regressing out residual motion effects based on the estimated motion variables does not work because the variables are correlated with the stimulus. They propose an additional correction based on modeling voxel-based experimental effects with and without residual motion-regression correction and show that their correction significantly affects the activation pattern for schizophrenics versus controls.

Other Issues

There are a number of other trends at this preprocessing stage: 1) using the pulse sequence to monitor rigid-body motion [54], 2) developing monitoring techniques as summary measures for QC [55] or to acquire the rigid-body transformations through external monitoring [56], and 3) volume-by-volume motion correction in real time [57].

Quality Control 2

At a minimum, plots of the six rigid body motion variables should be examined to determine the estimated range of motion in the data. It is also helpful to rerun the registration algorithm on the motion-corrected data and to compare these new movement variables from the motion-corrected data with the original estimates. Many centers use a rule-of-thumb that suggests that data should be discarded if estimated motion is more than 1–2 mm, although this typically depends on the scarcity of the data and the difficulty of replacing it. Such QC bounds may be applied directly to the estimated rigid-body motion plots or to calculated mean-voxel or maximum-voxel displacement values [35]. In addition, the motion variables should be correlated against the stimulus paradigm to see if there is strong evidence of stimulus-coupled motion. If large correlation coefficients are obtained, it may be necessary to consider additional processing steps such as those outlined by Bullmore et al. [53], although a more robust strategy would be to always evaluate SPI differences with and without residual motion regression. PICA and PCA are also useful to compare the change in components before and after movement correction, i.e., movement correction should increase the percent variance accounted for by stimulus-related components.

Physiological Noise Correction

Physiological noise in fMRI studies arises from cardiac-linked brain pulsations (~0.8 – 1.3 Hz), and bulk susceptibility variations of the chest with respiration that propagate as small magnetic field variations in the brain (0.1 – 0.3 Hz). Kruger and Glover [58] have demonstrated that at 3.0 T, these physiological noise sources are greater than those from system and thermal noise and significantly greater in gray matter than white matter. Provided the TR is fast enough to capture the unaliased cardiac effects (i.e., <500 ms), these noise sources may all be seen in the power spectrum and can be removed by notch filtering [59] or the more sophisticated multitaper techniques of Mitra and Pesaran [60]. The multitaper techniques

have also been used to reveal further intrinsic vascular pulsations at about 0.1 Hz [61]. However, most fMRI studies are conducted with TR $\geq$2 s, causing the cardiac noise to be aliased onto lower frequencies.

Many other approaches have been proposed for suppressing the cardiac and respiratory noise, and they fall into four groups: two based on *k*-space/image measurements with application of the estimated physiological correction variables in *k*-space [62], [63] or the reconstructed images [64], [65] and two using external monitoring of the cardiac and respiratory cycles to correct physiological effects in *k*-space [66] or in the reconstructed images [67]. These last two approaches using external physiological measurements are dubbed retrospective *k*-space correction (RETROKCOR) and retrospective image space correction (RETROICOR), respectively. Glover et al. [67] found that RETROICOR reduced localized physiological noise to a greater extent than RETROKCOR because the image-based technique is not limited, by adequate signal-to-noise, to *k*-space regions near the origin, i.e., higher spatial frequencies are smoothed out by RETROKCOR. Chuang and Chen [65] compared their image-based physiological artifacts estimation and correction technique (IMPACT) with RETROKCOR and found that IMPACT performed almost as well, suggesting that external physiological monitoring might not be necessary. However, the testing reported in the literature for these algorithms is quite limited, and none of the comparative testing reports appear to have considered interaction with head motion. A *k*-space correction occurs before motion correction, while an image-based correction may be performed before or after motion correction. Results for the comparison of these approaches are presumably dependent on the degree of image movement and on the particular subjects and populations being studied. However, much of the testing was performed with only a few slices of brain because of the temporal sampling requirement for capturing the cardiac cycle. In the most realistic dataset used by Glover (TR = 1,000 ms, 12 slices, 5-mm thick), which is still considerably faster than typical whole-brain fMRI studies with TR $\geq$ 2 s, subjects performed no task. In a more realistic whole-brain study with TR = 5 s and subjects performing a simple motor task, Tegeler et al. [9] demonstrated that RETROKCOR significantly improved run-to-run reliability for simple multivariate but not for univariate analyses in a few young subjects. A simple, block-design, finger-opposition task was scanned during three identical runs on a 4T scanner. The reliability of activations was measured across the runs as a function of the data analysis model [i.e., univariate *t*-test or multivariate Fisher's Linear Discriminant Analysis (FLDA)], with and without physiological corrections for respiration and respiration and cardiac image noise. Physiological corrections were performed using the externally monitored respiration and cardiac signals [66], and reliability was measured as the fraction of all the voxels activated across all three runs that appeared in only one run, appeared in any two runs, and appeared in all three runs. Respiration has a large impact on the fraction of highly reliable voxels for FLDA but not for the *t*-test. (Figure 5). There appears to be no data describing the relative importance of different types of physiological corrections in real, whole-brain imaging of populations susceptible to movement-related confounds.

Brain-Nonbrain Segmentation

Segmentation and classification of MRI brain voxels, MRI-MRI warping, and some cost functions used with fMRI-MRI alignment all perform better if the nonbrain tissues (e.g., scalp and skull) are identified and removed from MRI images. Many of the existing algorithms have recently been compared by Boesen et al. [68], Rex et al. [69], and Segonne et al. [70]. Boesen found that the consensus algorithm (Minneapolis Consensus Strip, McStrip) developed by Rehm et al. [71] performed better than any individual algorithm, and this finding was echoed by Rex but by using a different consensus approach based on training datasets and somewhat different algorithms. Segonne used a new hybrid approach with a watershed algorithm, which produces a segmented brain volume, to initialize a deformable surface model that integrates geometrical and atlas-based information developed on a sphere. Rehm provides a summary of validation results for 11 algorithms tested between 1997 and 2002, and Boesen

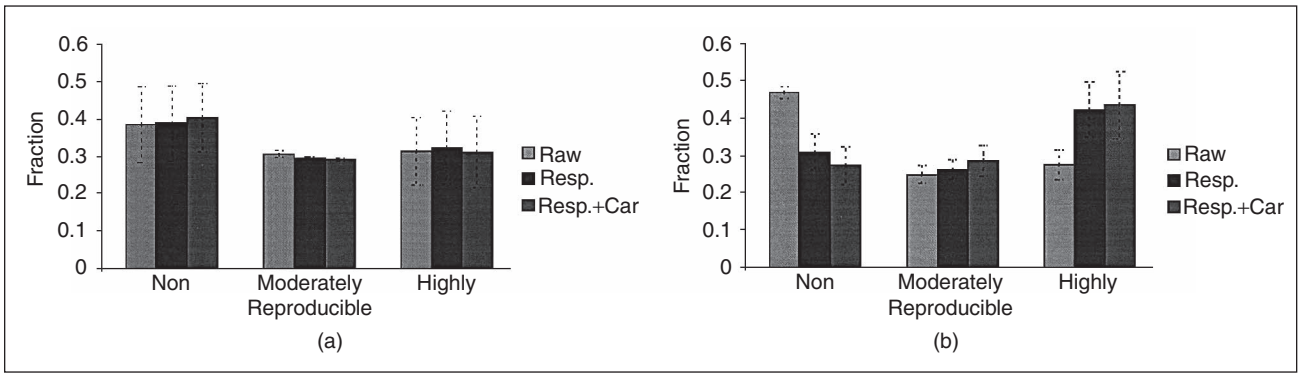

Fig. 5. A simple, block-design, finger-opposition task scanned during three identical runs on a 4 T scanner. The reliability of activations was measured across the runs as a function of the data analysis model (a) univariate *t*-test or (b) multivariate Fisher's linear discriminant analysis (FLDA), with and without physiological corrections for respiration (resp.) and respiration and cardiac (car.) image noise. Physiological corrections were performed using externally monitored respiration and cardiac signals (66), and reliability was measured as the fraction of all the voxels activated across all three runs that appeared in only one run (non), appeared in any two runs (moderately reproducible), and appeared in all three runs (highly). Respiration has a large impact on the fraction of highly reliable voxels for FLDA, but not for the *t*-test. Image is reproduced from (9) (*Human Brain Mapping*, Wiley and Sons, 1999).

demonstrates that McStrip consistently outperformed SPM2's brain extraction algorithm [72], FSL's Brain Extraction Tool (BET) [73], and the Brain Surface Extractor (BSE) [74]. Rex demonstrates that his Brain Extraction Meta-Algorithm (BEMA) outperforms BET, BSE, AFNI's 3dIntracranial program [75], and FreeSurfer's MRI Watershed [76]. Segonne compared the hybrid algorithm with BET, BSE, Freesurfer's strip skull [76], and Hahn and Peitgen's [77] watershed algorithm and found that it performed best across a variety of metrics with few misclassified voxels. All of the above algorithms aim to extract an accurate brain-surface boundary, generally defined as the outer pial surface of the cortex. However, Kovacevic et al. [78] developed an algorithm with the goal of extracting the brain and surrounding cerebrospinal fluid (CSF) in the subarachnoid space for studying longitudinal structural changes. This alternate extraction target is potentially less useful for fMRI preprocessing, although this hypothesis has yet to be tested.

Intensity Bias Correction

Structural MRI volumes tend to have nonuniform tissue intensities, particularly at higher fields, as a result of the nonuniform sensitivity profiles of the radio frequency coils used to transmit and receive the signal in MRI scanners [6]. Significant nonuniformities may bias the results of algorithms that utilize image intensities (e.g., segmentation and registration) in their cost functions and should be corrected. A variety of algorithms have been developed for correcting intensity nonuniformities and new approaches are being continually developed [79], with an emphasis on correction in the presence of atrophy from dementia and aging [80] and longitudinal studies of structural changes over time [81]. A quantitative comparison of six techniques was performed by Arnold et al. [82] in which two locally adaptive algorithms provided the most reliable correction: nonparametric, nonuniform intensity normalization (N3) [83], and bias field corrector (BFC) [74]. Shattuck compared N3 and BFC and found that BFC outperformed N3 on phantom data. There are many algorithms, such as BFC, that are incorporated into tissue segmentation and classification packages because bias correction is a necessary preprocessing step for (and interacts with) tissue classification [84]. As a result, there are algorithms for correcting intensity nonuniformities that are part of tissue segmentation and classification packages that have not been compared, even across the major software packages, e.g., from within FMRIB's automated segmentation tool (FAST) [85].

Quality Control 3

At a minimum, the results of brain-nonbrain extraction and intensity bias correction must be visually assessed and repeated, possibly with a different algorithm/cost function, if clearly unsatisfactory. Training examples of satisfactory and unsatisfactory results should be available for visual comparisons. Quantitative metrics for assessing these preprocessing steps are urgently needed.

Registration II: fMRI-MRI

After physiological correction, following the green preprocessing path in Figure 2 (see "Preprocessing Paths"), the registered fMRI time series is aligned with a high-resolution MRI (fMRI-to-MRI) that may have undergone brain-nonbrain segmentation and intensity bias correction. The MRI is usually collected during the same scanning session and has much less distortion and better tissue contrast than EPI or SI fMRI scans. It is collected 1) to visualize single-subject task activations overlaid on each individual subject's anatomical locations; 2) to align the fMRI scans to a common coordinate space defined by a structural template, combining measured fMRI-to-MRI and nonlinear MRI-template registration transformations (i.e., warps), before data analysis; and 3) to warp single-subject data-analysis results to the template's coordinate space for higher-level group analysis.

After Registration I, a mean fMRI scan is usually calculated and used as the target to calculate the fMRI-MRI transformation variables. The MRI is usually resampled into the fMRI space because the higher-resolution image will suffer less distortion after interpolation. At least an affine (12 parameter) fMRI-MRI transformation should be calculated to allow for scaling between the fMRI and MRI images, and with a robust cost function (e.g., mutual information), this will not require the brain to be segmented from the nonbrain portions of the MRI before alignment, as described in the following. If using AIR's cost functions [30], which do not include mutual information, the best results are achieved if the brain is segmented from the nonbrain tissue before registration. Note that while mutual information and correlation cost functions have been proposed as optimal for fMRI-fMRI alignment, to my knowledge there is no paper comparing the major packages' performance for the fMRI-MRI registration problem. However, Cusack et al. [49] show that if EPI images are corrected for magnetic field inhomogeneities, the values of optimal mutual information obtained between fMRI and MRI using SPM99 are significantly improved.

As one approach to solving the individual slice movement problem, the so-called map-slice-to-volume (MSV) technique, aligns individual fMRI slices to each subject's MRI acquired during the same scanning session. This approach was successfully demonstrated by Kim et al. [44] and has recently been extended to include joint-estimation of each slice's orientation under smoothness constraints on the subject's movement trajectory over time, i.e., joint MSV. In addition, the MSV approach has been combined with a concurrent iterative image reconstruction incorporating correction for field-inhomogeneities [47]. To my knowledge, this is the first combined solution that corrects for the interactions between slice timing, field-inhomogeneities, and fMRI-to-fMRI image registration, while also achieving fMRI-MRI alignment.

Different Preprocessing Paths

There are two different preprocessing paths: 1) an all-in-one analysis (green path, Figure 2), in which raw data is completely preprocessed before data analysis, as is typical in SPM, and 2) a summary statistic approach (red path, Figure 2), with individual-subject preprocessing and data analysis, without applying Registration II and III to the fMRI data, followed by registration of the resulting SPI's for further higher-level data analysis, as is typical in FSL [84].

The structural MRI preprocessing is used by both approaches (blue path, Figure 2). The all-in-one analysis includes (green path), Registration I, physiological correction, Registration II and III, and spatial filtering through to a final SPI for both individual subjects and groups. If planning a group analysis, the motion- and physiology-corrected fMRIs would typically be passed straight to Registration III, together with the fMRI-MRI

transformation, and bias-corrected and brain-nonbrain segmented MRIs; in order to calculate the fMRI-MRI transformation, the averaged fMRI would have previously been passed to Registration II. In Registration III, these elements are combined to create a set of subjects' fMRI scans aligned to a common coordinate system (see Registration III), and these scans are passed on to spatial smoothing.

The summary statistics approach (red path) includes Registration I, physiological correction, and spatial filtering through to an SPI for each subject. Transformations from Registration II are then applied to individual SPIs so that they may be displayed on their subjects' MRIs, and by combining transformations from Registration II and III, SPIs are resampled into a common space to perform a higher-level group analysis. Note that the SPIs generated for passing to a higher-level analysis are different from those typically output for thresholding to detect brain activations.

There are a number of other possible combinations of the above steps: 1) performing physiological correction before Registration I or after Registration III, 2) combining transformations from Registration I, II, and III into a single resampling step of the raw data into a common space, which minimizes interpolation smoothing effects, and 3) collecting both low-resolution, structural EPI/SI scans that match the low-resolution fMRI scans and high-resolution MRIs in the same subject; the fMRI scans are aligned to the low-resolution structural EPI/spiral scans, which are warped to the high-resolution MRI to complete Registration II [86].

Quality Control 4

The fMRI-MRI alignment should be visually assessed. Training examples of satisfactory and unsatisfactory results should be available for visual comparisons. Quantitative metrics for assessing this preprocessing step are urgently needed.

Registration III: MRI-Template

There are two broad categories of algorithm based on cost functions using surface-based landmarks or voxel-based intensities. I will primarily focus on intensity-based algorithms as they are the most widely used and are readily available in the major software packages. I will also briefly discuss combined landmark and intensity-based schemes for image volumes and their recent extension to surface-based registration approaches. Following Ardekani et al. [87], intensity-based matching approaches may be categorized into 1) piecewise-linear (Tailarach transformation [88] in AFNI [89]) and affine transformations (AIR [90]; FLIRT in FSL, [25]; SPM [27]) with relatively few parameters, 2) nonlinear transformations (i.e., warps) with hundreds of degrees of freedom (SPM, [91]; AIR, [90]), and 3) highly flexible warps with thousands of degrees of freedom [92]–[99]. Note that the higher-order algorithms in 2) and 3) above are typically initialized with a low dimensional, usually affine, transformation from category 1). Jenkinson et al. [25] compared the reliability of such affine registrations for between-subject registration, including AIR and SPM99, and found that the robustness of the optimization procedure implemented in FSL's FLIRT had some advantages in generating consistent transformations.

Structural templates are used to define a common coordinate system that may be used to locate similar regions in the brains of multiple subjects. The most important is the Talairach system based on a single older brain [88]. Most groups have moved away from the original piecewise linear transformation defined by Talairach for each individual MRI in favor of a population standard defined by many MRIs. There are a number of such standard templates developed by the Montreal Neurological Institute (MNI) and the International Consortium for Brain Mapping. These templates are introduced and their relationship to Talairach coordinates is described by Brett et al. [100], with additional details and web links at http://www.mrccbu.cam.ac.uk/Imaging/Common/mnispace.shtml. As a general rule, the template should match the data being aligned as closely as possible. Therefore, leaving out Registration II completely and nonlinearly warping the average of the movement-corrected fMRI time series directly to a matched template (e.g., EPI fMRI scans to an EPI structural template) may minimize the effects of geometric distortion from magnetic field inhomogeneities. However, this approach may be less robust than the use of MRIs through Registration II, due to the lower quality and lower resolution of EPIs compared to MRIs [49]. I would expect this result to depend on scanner-dependent image artifacts and whether or not the template is generic (e.g., in one of the packages) or specifically built for a particular scanner or even group analysis as outlined in the following.

An alternative is to develop a group-specific template that minimizes the differences between itself and all scans for the group of subjects being analyzed. There have been several proposals for developing such optimal, group-specific templates [28], [101]–[104], but to date their use has not been compared with the standard MNI/ICBM population-based templates.

Hellier et al. [105] have shown that global metrics assessing the registration of structural MRIs improve with increasing degrees of freedom, but local measures of the match between cortical sulci are not significantly different between simple, six-parameter, rigid-body transformations and highly flexible warps. As a possible means of addressing this surprising result, a number of groups are exploring the use of intensity-based algorithms with additional anatomical constraints, such as major sulcal landmarks or automatically generated features [106]–[108]. Magnotta et al. [109] have shown average increases in relative regional overlap with increases in both the degrees of freedom of an intensity-based algorithm and the amount of anatomical landmark information used. Recent developments include the use of structural landmarks to constrain warps across cortical surface representations on a sphere or a flattened representation [110], [111].

There are few comparisons of the impact of MRI-based anatomical warps on functional SPIs from groups. Using positron emission tomography (PET) and MRI scans, Kjems et al. [92] demonstrated increased functional activation signals in a group analysis with a multiscale, intersubject warp evaluated with different cost functions versus affine transformations. They provided preliminary evidence that improved structural warps do not necessarily translate into improved group functional analysis results. Using a similar metric, Dinov et al. [112] compared PET activation resulting from seven- and 12–parameter affine and second- and fifth-order polynomial warps from AIR 3.0. They found an optimal activation peak for a 12-parameter affine transformation. Crivello et al. [113] also used PET and MRI scans with an affine and three warping algorithms: affine and the fifth-order polynomial in AIR 3.0, the basis-function warp in SPM96 [26], and the flexible, multigrid approach of Schormann and Zilles [114].

They found that the expected improvement in structural alignment with increasing degrees of freedom was matched by only limited improvements in group functional analysis results, which were dependent on the amount of spatial smoothing applied to the PET datasets. These findings agree with the recent fMRI group-analysis evaluations by Strother et al. [1]. They show that there is a small improvement in global functional performance metrics with third- to seventh-order polynomial warps compared to 12-parameter affine registration using AIR 5.03 [90] and that this improvement is a function of fMRI spatial smoothing. However, Ardekani et al. [87] compared AFNI's Talairach registration with SPM99's warps and a highly flexible in-house algorithm (ART) and found a significant improvement in group-analysis results using ART. One difference is that Ardekani aligned the individual subject's SPIs rather than analyzing the aligned, raw fMRI images as in Strother et al. Evidence seems to be accumulating that highly accurate structural warps are not necessary for functional group analysis, but there is much more work required to understand the structure-function relationship in young, normal brains, let alone that found in aging and abnormal brains.

Spatial Filtering

The importance of measuring and manipulating the spatial correlations of functional neuroimages has been extensively studied [115]–[117], and the dramatic impact a little spatial smoothing may have on signal detection power in fMRI should be carefully considered when specifying a preprocessing pipeline [1], [118]–[120]. In Figure 2, the spatial filtering step is placed before temporal filtering and unwanted-effects' regressors to improve variable estimates within these steps. Spatial filtering follows physiological corrections in Figure 2 to avoid smearing large, local physiological noise sources into surrounding tissue. Simulation results from Skudlarski et al. [118] illustrate the increased signal-detection power obtained by smoothing either the raw data or SPIs. Recent results support the Gaussian smoothing kernel defaults of 1–2 voxels FWHM in the major packages and indicate that spatial smoothing is one of the most influential preprocessing choices when compared with other steps [1], [121]. Shaw et al. [11] have shown that optimal spatial smoothing settings vary from subject to subject, reinforcing the potential advantages of locally adaptive smoothing approaches [122]–[124] or combining results across multiple spatial smoothing scales [115], [117], [118]. Wink and Roerdink [125] compared spatial wavelet denoising with Gaussian smoothing and reported that, "wavelet denoising methods that introduce relatively little smoothness are generally preferable over Gaussian smoothing for denoising fMRI time series."

An analysis approach related to spatial smoothing is the definition and extraction of averaged voxel values in anatomically or functionally defined regions- or volumes-of-interest (ROIs or VOIs). This approach to data extraction and analysis was developed and widely used in the PET community in the 1980s and early 1990s before the development of voxel-based analysis techniques [126] and is becoming more common for fMRI analysis [127]. I expect the interaction of preprocessing steps with the definition and analysis of ROIs to become an important topic, particularly for network analysis using structural equation models, which require a small number of stable time series to be extracted from predefined regions of the brain [10].

Intensity Normalization

In fMRI experiments there is additional scan-to-scan variance at very low spatial frequencies that cannot be readily accounted for by the experimental stimulus. A variety of global normalization approaches have developed to model and control for this variation, originally in PET [128], [129] and, based on these, more recently in fMRI [130]–[132]. Gavrilescu et al. compared five approaches: grand-mean session scaling (GMS), proportional scaling (PS), ANCOVA, iterative masking of activation areas [129], and orthogonalization of the global estimate [131]. They propose a flowchart that chooses a normalization technique based on the comparison of SPIs generated using GMS and PS and the measurement of the correlation between scan means and GLM covariates of interest, such as the stimulus reference function. Skudlarski et al. [118] found little benefit for three approaches to PS using simulations but noted that the appearance of large global fluctuations is very subject dependent, and some subjects may benefit from intensity normalization.

These intensity normalization techniques interact with the temporal and spatial filtering approaches because 1) to the extent that global variations are low- or high-temporal frequency phenomena, they will tend to be removed by low- and high-pass temporal filters, and 2) spatial smoothing tends to increase the correlation between voxel time series and global signals [130]. Using a spatio-temporal basis representation such as wavelets [133], it may be possible to identify and eliminate large-scale spatial (i.e., global) components, but this use of spatio-temporal basis approaches has not been systematically investigated. It is likely that in the future adaptive, spatio-temporal data analysis models will unify questions of spatial and temporal filtering and intensity normalization within a single step [124].

Temporal Filtering

Previously, I addressed frequency domain filtering in the context of physiological noise correction. In the following, I focus on low-pass filtering for denoising, temporal smoothing for control and estimation of temporal autocorrelations (i.e., precoloring), and high-pass filtering for control of low-frequency noise trends. Low-pass filtering, or temporal smoothing, applied to individual voxel time series has been found to degrade signal detection power in block designs [118] and degrades parameter estimation efficiency even more for single-event designs [134]. Precoloring was advocated within the SPM package before SPM2 [135] to stabilize and obtain unbiased inferential estimates. However, it is now clear that prewhitening—removal of temporal autocorrelation to whiten the residual noise—is feasible and should be used to improve parameter estimation efficiency for inferential significance testing of GLMs applied to single-voxel time series [84], [136]. However, if using nonparametric tests that control for temporal autocorrelation [133], it is unclear if either prewhitening or precoloring should be used. A related technique to low-pass filtering is the removal of basis components representing higher-order, spatio-temporal frequencies, i.e., basis function denoising. For example, Strother et al. [1] have demonstrated that when using a linear discriminant built on a PCA basis more than 90% of the low-variance principal components should be discarded to maximize the linear discriminant performance. Thomas et al. [21] compared denoising of fMRI time series with ICA and PCA and found that both

Many researchers using fMRI implicitly assume that the exact choices of testing environment and performance metrics do not matter.

improved the ability to detect a BOLD signal change in the presence of physiological and scanner noise; ICA was best for isolation and removal of structured noise, and PCA was superior for isolation and removal of random noise.

The noise power in the majority of fMRI time series occurs at low frequencies ($f < 0.05$ Hz) with an approximately 1/f envelope. The exact causes of this noise structure remain unclear, although it is probably some mixture of scanner drift [137], aliased cardiac physiological noise, and respiratory physiological noise, such as that due to resting fluctuations in arterial CO2 [138]. Whatever the cause, it is not uncommon to find adjacent voxels with quite different low-frequency trends, e.g., positive and negative slopes. Numerous reports have demonstrated the importance of removing these trends to improve signal power [118], [121], but care must be taken to not attenuate the signal, which in block designs often has a relatively low fundamental frequency of around 0.02 Hz (i.e., baseline and activation epochs of 25 s each). Removal may be effected by explicit low-pass filtering in the frequency domain—Skudlarski found a Butterworth filter with a cutoff of 0.35 of the stimulus frequency performed best [118]—or by regressing out linear and higher-order polynomial trends as unwanted effects' regressors in a GLM. Alternatively, SPM provides a user-chosen set of cosine basis functions that are removed by GLM regression from each voxel's time series [139]. Optimal preprocessing pipeline performance has been demonstrated to require adaptive tuning of the number of detrending, cosine basis functions in groups [1], and individual subjects [140]. Tanabe et al. [141] tested five trend removal techniques and found that a cubic spline produced the largest number of significant voxels among linear, quadratic, cubic, and wavelet detrending approaches. In addition, while 42% of the voxels tested required no detrending, the techniques producing the smallest p-values in each of the remaining voxels were cubic splines in about 28% and quadratic and wavelets in about 15% each. Finally, the FSL package has adopted a quite different nonlinear trend removal approach using a running lines smoother [142]. Marchini et al. [142] also note the need to deal with isolated spikes by reducing large deviations from the mean to within some specified limits. In summary, there is no consensus as to which high-pass filtering or detrending approach to use with growing evidence that the technique(s) should be adapted to the subject or even the voxel being analyzed.

Quality Control 5

Except for unwanted-effects regressors (see the following) that are often removed as part of a GLM data analysis, the preprocessing is now complete. Whether or not it is being used as a data analysis tool, a PICA or PCA should be run at this point and compared with those run at QC1 and QC2 to examine the reductions in variance and removal of artifactual components by the preprocessing pipeline. Any remaining artifactual components may also be identified.

Unwanted-Effects Regressors

These include removal by regression within a GLM of residual-movement artifacts, low-frequency trends, global intensity variations (by ANCOVA), and even physiological time series, all of which have been discussed above. In Figure 2 the step is half in the shaded data analysis box because it may be combined with a GLM data analysis. It is half out of the shaded box because a GLM to remove unwanted effects may be run as an independent preprocessing step before another data analysis approach is used. This is the approach taken in Strother et al. [1] where the fMRI data from spatial filtering undergoes a PCA followed by denoising—dropping 93% of the principal components—and detrending of the remaining eigenvector time series in a separate GLM step before applying a linear discriminant analysis.

Conclusions

With the present technology-focused approach to testing individual preprocessing steps, each paper tends to use a different testing environment and performance metrics. This would be less important if they were all testing exactly the same combinations of algorithms and software implementations, but they are not, leaving the generality of the reported testing results largely unknown. Based on the existing literature, it is impossible to make conclusive statements about the optimal algorithms and software implementations for any single preprocessing step, let alone entire pipelines. Many researchers using fMRI implicitly assume, or perhaps hope, that the exact choices do not matter, but at present that is a largely untested scientific hypothesis! As outlined in Strother et al. [1], there is growing evidence from both real data and simulations that by applying a new processing pipeline to a raw dataset, significantly modified spatial activation patterns may be obtained as a result of changing/optimizing preprocessing techniques and/or the data analysis approach.

I believe that the present focus on the technological testing of preprocessing steps should be balanced by approaches that test the whole pipeline. This should include all interactions measured using metrics that are closely linked to the research and diagnostic questions to be addressed at the end of the processing pipeline (Figure 1). The goal is to avoid single expedient or default pipelines by developing a framework capable of potentially testing thousands of possible pipeline implementations per dataset. To achieve this goal, we are depending on recent developments in software tools for managing neuroimaging workflows [143]–[145]. These flexible frameworks for assembling and testing heterogeneous pipelines will be used to determine the most important steps,

algorithms, software implementations, and their parameters for any given dataset. Here, the pipeline is thought of as a flexible meta-model, defined by all of the choices and parameters involved in the acquisition and reconstruction, preprocessing, and data analysis steps. These meta-model choices and parameters should be optimized for the task at hand. We have made a start on such a testing program using the resampling framework reported in [1], [11], [121], [146], [147]. In addition to the traditional, technologically-oriented testing of new algorithms and their associated software tools, I believe that the functional neuroimaging field should now enter a new phase of testing in which the optimization of complete, heterogeneous processing pipelines is emphasized.

Acknowledgments

I wish to thank my colleagues at the Rotman Research Institute, University of Toronto, and the International Neuroimaging Consortium, University of Minnesota, for helpful discussions. Helpful suggestions from two anonymous reviewers also significantly improved this article. This work was partly supported by the National Institutes of Health Human Brain Project P20 Grant EB02013.

Stephen C. Strother is a senior scientist in the Rotman Research Institute at Baycrest and a professor of medical biophysics at the University of Toronto, Canada. He specializes in data analysis and optimization of positron emission tomography (PET) and fMRI neuroimaging techniques for both research and clinical applications. He is best known for championing the use of multivariate analysis and statistical learning techniques and using these to develop quantitative performance metrics in functional neuroimaging. In 2001, he cofounded Predictek, LLC, a Chicago-based consulting company that focuses on predictive modeling with neuroimages for the pharmaceutical and vision care industries. Since 2002, he has chaired the Data Format Working Group, an international standards committee for fMRI data exchange set up under the Neuroimaging Informatics Technology Initiative at the National Institutes of Health in the United States. He received his Ph.D. in electrical engineering from McGill University in Montreal, was a postdoctoral fellow at Memorial Sloan-Kettering Cancer Center, New York, and was then based at the VA Medical Center and the University of Minnesota in Minneapolis until he moved to Toronto in 2004.

Address for Correspondence: Stephen C. Strother, Ph.D., Senior Scientist, Rotman Institute-Baycrest Centre, 3560 Bathurst Street, Toronto, ON, M6A-2E1 Canada. Phone: +1 416 785 2500 ext. 2956. E-mail: sstrother@rotman-baycrest.on.ca.

References

[1] S. Strother, S. La Conte, L. Kai Hansen, J. Anderson, J. Zhang, S. Pulapura, and D. Rottenberg, "Optimizing the fMRI data-processing pipeline using prediction and reproducibility performance metrics: I. A preliminary group analysis," *NeuroImage,* vol. 23, Suppl. 1, pp. S196–S207, 2004.

[2] V.D. Calhoun, T. Adali, G.D. Pearlson, P.C. van Zijl, and J.J. Pekar, "Independent component analysis of fMRI data in the complex domain," *Magn. Reson. Med.*, vol. 48, no. 1, pp. 180–192, 2002.

[3] D. Rowe, "Modeling both the magnitude and phase of complex-valued fMRI data," *NeuroImage*, vol. 25, no. 4, pp. 1310–1324, May 1, 2005.

[4] D. Rowe, "Parameter estimation in the magnitude-only and complex-valued fMRI data models," *NeuroImage*, vol. 25, no. 4, pp. 1124–1132, 2005.

[5] R.S. Menon, "Postacquisition suppression of large-vessel BOLD signals in high-resolution fMRI," *Magn. Reson. Med.,* vol. 47, no. 1, pp. 1–9, 2002.

[6] S.A. Huettel, A.W. Song, and G. McCarthy, *Functional Magnetic Resonance Imaging*. Sunderland, MA: Sinauer Assoc., 2004.

[7] P. Jezzard, P.M. Matthews, and S.M. Smith, *Functional Magnetic Resonance Imaging: An Introduction to Methods,* 2001.

[8] A.W. Toga and J.C. Mazziotta, *Brain Mapping: The Methods*, 2nd ed. San Diego, CA: Academic, 2002.

[9] C. Tegeler, S.C. Strother, J.R. Anderson, and S.G. Kim, "Reproducibility of BOLD-based functional MRI obtained at 4 T," *Hum. Brain Map.*, vol. 7, no. 4, pp. 267–283, 1999.

[10] M. Gavrilescu, G.W. Stuart, A. Waites, G. Jackson, I.D. Svalbe, and G.F. Egan, "Changes in effective connectivity models in the presence of task-correlated motion: An fMRI study," *Hum. Brain Map.*, vol. 21, no. 2, pp. 49–63, 2004.

[11] M.E. Shaw, S.C. Strother, M. Gavrilescu, K. Podzebenko, A. Waites, J. Watson, J. Anderson, G. Jackson, and G. Egan, "Evaluating subject specific preprocessing choices in multisubject fMRI data sets using data-driven performance metrics," *NeuroImage*, vol. 19, no. 3, pp. 988–1001, 2003.

[12] J.S. Gati, R.S. Menon, K. Ugurbil, and B.K. Rutt, "Experimental determination of the BOLD field strength dependence in vessels and tissue," *Magn. Reson. Med.,* vol. 38, no. 2, pp. 296–302, 1997.

[13] G. Kruger, A. Kastrup, and G.H. Glover, "Neuroimaging at 1.5 T and 3.0 T: Comparison of oxygenation-sensitive magnetic resonance imaging," *Magn. Reson. Med.*, vol. 45, no. 4, pp. 595–604, 2001.

[14] S. Gold, B. Christian, S. Arndt, G. Zeien, T. Cizadlo, D.L. Johnson, M. Flaum, and N.C. Andreasen, "Functional MRI statistical software packages: A comparative analysis," *Hum. Brain Map.*, vol. 6, no. 2, pp. 73–84, 1998.

[15] J.D. Van Horn and M.S. Gazzaniga, "Opinion: Databasing fMRI studies towards a 'discovery science' of brain function," *Nat. Rev .Neurosci.*, vol. 3, no. 4, pp. 314–318, 2002.

[16] J.D. Van Horn, "Number of distributable data sets originally processed with SPM.," S.C. Strother, Ed., private communication, 2005.

[17] W.L. Luo and T.E. Nichols, "Diagnosis and exploration of massively univariate neuroimaging models," *NeuroImage*, vol. 19, no. 3, pp. 1014–1032, 2003.

[18] J.J. Sychra, P.A. Bandettini, N. Bhattacharya, and Q. Lin, "Synthetic images by subspace transforms. I. Principal components images and related filters," *Med. Phys.*, vol. 21, no. 2, pp. 193–201, 1994.

[19] M.J. McKeown, T.P. Jung, S. Makeig, G. Brown, S.S. Kindermann, T.W. Lee, and T.J. Sejnowski, "Spatially independent activity patterns in functional MRI data during the stroop color-naming task," *Proc. Nat. Acad. Sci. U.S.A.*, vol. 95, no. 3, pp. 803–810, 1998.

[20] C.F. Beckmann and S.M. Smith, "Probabilistic independent component analysis for functional magnetic resonance imaging," *IEEE Trans. Med. Imag.*, vol. 23, no. 2, pp. 137–152, 2004.

[21] C.G. Thomas, R.A. Harshman, and R.S. Menon, "Noise reduction in BOLD-based fMRI using component analysis," *NeuroImage*, vol. 17, no. 3, pp. 1521–1537, 2002.

[22] R.W. Cox, J. Ashburner, H. Breman, K. Fissell, C. Haselgrove, C.J. Holmes, J.L. Lancaster, D. Rex, S.M. Smith, J.B. Woodward, and S.C. Strother, "A (Sort of) new image data format standard," NIfTI-1, Budapest, Hungary, 2004.

[23] E. Seto, G. Sela, W.E. McIlroy, S.E. Black, W.R. Staines, M.J. Bronskill, A. R. McIntosh, and S.J. Graham, "Quantifying head motion associated with motor tasks used in fMRI," *NeuroImage*, vol. 14, no. 2, pp. 284–297, 2001.

[24] R.W. Cox and A. Jesmanowicz, "Real-time 3D image registration for functional MRI," *Magn. Reson. Med.*, vol. 42, no. 6, pp. 1014–1018, 1999.

[25] M. Jenkinson, P. Bannister, M. Brady, and S. Smith, "Improved optimization for the robust and accurate linear registration and motion correction of brain images," *NeuroImage*, vol. 17, no. 2, pp. 825–841, 2002.

[26] K.J. Friston, J. Ashburner, J.B. Poline, C.D. Frith, J.D. Heather, and R.S.J. Frackowiak, "Spatial registration and normalization of images," *Hum. Brain Map.,* vol. 3, no. 3, pp. 165–189, 1995.

[27] J. Ashburner and K. Friston, "Multimodal image coregistration and partitioning—a unified framework," *NeuroImage*, vol. 6, no. 3, pp. 209–217, 1997.

[28] R.P. Woods, M. Dapretto, N.L. Sicotte, A.W. Toga, and J.C. Mazziotta, "Creation and use of a Talairach-compatible atlas for accurate, automated, nonlinear intersubject registration, and analysis of functional imaging data," *Hum. Brain Map.*, vol. 8, no. 2-3, pp. 73–79, 1999.

[29] A. Jiang, D.N. Kennedy, J.R. Baker, R.M. Weisskoff, R.B.H. Tootell, R.P. Woods, R.R. Benson, K.K. Kwong, T.J. Brady, B.R. Rosen, and J.W. Belliveau, "Motion detection and correction in functional MR imaging," *Hum. Brain Map.*, vol. 3, no. 3, pp. 224–235, 1995.

[30] R.P. Woods, S.T. Grafton, C.J. Holmes, S.R. Cherry, and J.C. Mazziotta, "Automated image registration: I. General methods and intrasubject, intramodality validation," *J. Comput. Assisted Tomography*, vol. 22, no. 1, pp. 139–152, 1998.

[31] B.A. Ardekani, A.H. Bachman, and J.A. Helpern, "A quantitative comparison of motion detection algorithms in fMRI," *Magn. Reson. Imag.*, vol. 19, no. 7, pp. 959–963, 2001.

[32] V.L. Morgan, D.R. Pickens, S.L. Hartmann, and R.R. Price, "Comparison of functional MRI image realignment tools using a computer-generated phantom," *Magn. Reson. Med.*, vol. 46, no. 3, pp. 510–514, 2001.

[33] L. Freire and M. Jenkinson, "A gradient-informed robust motion correction method for FMRI," in *Biomedical Image Registration*, vol. 2717, *Lecture Notes Comput. Sci.*, pp. 377–386, 2003.

[34] L. Freire, J. Orchard, M. Jenkinson, and J.F. Mangin, "Reducing activation-related bias in FMRI registration," in *Medical Imaging and Augmented Reality, Proceedings*, vol. 3150, *Lecture Notes in Computer Science*, pp. 278–285, 2004.

[35] S.C. Strother, J.R. Anderson, X.L. Xu, J.S. Liow, D.C. Bonar, and D.A. Rottenberg, "Quantitative comparisons of image registration techniques based on

high-resolution MRI of the brain," *J. Comput. Assisted Tomography*, vol. 18, no. 6, pp. 954–962, 1994.
[36] J. West, J.M. Fitzpatrick, M.Y. Wang, B.M. Dawant, C.R. Maurer, Jr., R.M. Kessler, R.J. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P.A. van den Elsen, S. Napel, T.S. Sumanaweera, B. Harkness, P.F. Hemler, D.L. Hill, D.J. Hawkes, C. Studholme, J.B. Maintz, M.A. Viergever, G. Malandain, R.P. Woods, and et al., "Comparison and evaluation of retrospective intermodality brain image registration techniques," *J. Comput. Assisted Tomography*, vol. 21, no. 4, pp. 554–566, 1997.
[37] L. Freire, A. Roche, and J.F. Mangin, "What is the best similarity measure for motion correction in fMRI time series?," *IEEE Trans. Med. Imag.*, vol. 21, no. 5, pp. 470–484, 2002.
[38] L. Freire and J.F. Mangin, "Motion correction algorithms may create spurious brain activations in the absence of subject motion," *NeuroImage*, vol. 14, no. 3, pp. 709–722, 2001.
[39] W.F. Eddy, M. Fitzgerald, and D.C. Noll, "Improved image registration by using Fourier interpolation," *Magn. Reson. Med.*, vol. 36, no. 6, pp. 923–931, 1996.
[40] J.V. Hajnal, N. Saeed, E.J. Soar, A. Oatridge, I.R. Young, and G.M. Bydder, "A registration and interpolation procedure for subvoxel matching of serially acquired MR images," *J. Comput. Assisted Tomography*, vol. 19, no. 2, pp. 289–296, 1995.
[41] T.M. Lehmann, C. Gonner, and K. Spitzer, "Addendum: B-spline interpolation in medical image processing," *IEEE Trans. Med. Imag.*, vol. 20, no. 7, pp. 660–665, 2001.
[42] T.M. Lehmann, C. Gonner, and K. Spitzer, "Survey: Interpolation methods in medical image processing," *IEEE Trans. Med. Imag.*, vol. 18, no. 11, pp. 1049–1075, 1999.
[43] S. Grootoonk, C. Hutton, J. Ashburner, A.M. Howseman, O. Josephs, G. Rees, K.J. Friston, and R. Turner, "Characterization and correction of interpolation effects in the realignment of fMRI time series," *NeuroImage*, vol. 11, no. 1, pp. 49–57, 2000.
[44] B. Kim, J.L. Boes, P.H. Bland, T.L. Chenevert, and C.R. Meyer, "Motion correction in fMRI via registration of individual slices into an anatomical volume," *Magn. Reson. Med.*, vol. 41, no. 5, pp. 964–972, 1999.
[45] P.R. Bannister, J.M. Brady, and M. Jenkinson, "TIGER—A new model for spatio-temporal realignment of FMRI data," in *Computer Vision and Mathematical Methods in Medical and Biomedical Image Analysis*, vol. 3117, *Lecture Notes in Computer Science*, pp. 292–303, 2004.
[46] P. Jezzard and R.S. Balaban, "Correction for geometric distortion in echo planar images from B0 field variations," *Magn. Reson. Med.*, vol. 34, no. 1, pp. 65–73, 1995.
[47] B.P. Sutton, D.C. Noll, and J.A. Fessler, "Fast, iterative image reconstruction for MRI in the presence of field inhomogeneities," *IEEE Trans. Med. Imag.*, vol. 22, no. 2, pp. 178–188, 2003.
[48] C. Hutton, A. Bork, O. Josephs, R. Deichmann, J. Ashburner, and R. Turner, "Image distortion correction in fMRI: A quantitative evaluation," NeuroImage, vol. 16, no. 1, pp. 217–240, 2002.
[49] R. Cusack, M. Brett, and K. Osswald, "An evaluation of the use of magnetic field maps to undistort echo-planar images," *NeuroImage*, vol. 18, no. 1, pp. 127–142, 2003.
[50] M.A. Elliott, E.E. Gualtieri, J. Hulvershorn, J.D. Ragland, and R. Gur, "The effects of geometric distortion correction on motion realignment in fMRI," *Acad. Radiology*, vol. 11, no. 9, pp. 1005–1010, 2004.
[51] J.L.R. Andersson, C. Hutton, J. Ashburner, R. Turner, and K. Friston, "Modeling geometric deformations in EPI time series," *NeuroImage*, vol. 13, no. 5, pp. 903–919, 2001.
[52] K.J. Friston, S. Howard, R.S.J. Frackowiak, and R. Turner, "Movement-related effects in fMRI time-series," *Magn. Reson. Med.*, vol. 35, no. 3, pp. 346–355, 1996.
[53] E.T. Bullmore, M.J. Brammer, S. Rabe-Hesketh, V.A. Curtis, R.G. Morris, S.C. Williams, T. Sharma, and P.K. McGuire, "Methods for diagnosis and treatment of stimulus-correlated motion in generic brain activation studies using fMRI," *Hum. Brain Map.*, vol. 7, no. 1, pp. 38–48, 1999.
[54] E.B. Welch, A. Manduca, R.C. Grimm, H.A. Ward, and C.R. Jack, "Spherical navigator echoes for full 3D rigid body motion measurement in MRI," *Magn. Reson. Med.*, vol. 47, no. 1, pp. 32–41, 2002.
[55] E.C. Caparelli, D. Tomasi, S. Arnold, L. Chang, and T. Ernst, "K-space based summary motion detection for functional magnetic resonance imaging," *NeuroImage*, vol. 20, no. 2, pp. 1411–1418, 2003.
[56] M. Tremblay, F. Tam, S.J. Graham, "Retrospective coregistration of functional magnetic resonance imaging data Using External Monitoring," *Magn. Reson. Med.*, vol. 53, no.1, 2005.
[57] K. Mathiak and S. Posse, "Evaluation of motion and realignment for functional magnetic resonance imaging in real time," *Magn. Reson. Med.*, vol. 45, no. 1, pp. 167–171, 2001.
[58] G. Kruger and G.H. Glover, "Physiological noise in oxygenation-sensitive magnetic resonance imaging," *Magn. Reson. Med.*, vol. 46, no. 4, pp. 631–637, 2001.
[59] B. Biswal, A.E. DeYoe, and J.S. Hyde, "Reduction of physiological fluctuations in fMRI using digital filters," *Magn. Reson. Med.*, vol. 35, no. 1, pp. 107–13, 1996.
[60] P.P. Mitra and B. Pesaran, "Analysis of dynamic brain imaging data," *Biophys. J.*, vol. 76, no. 2, pp. 691–708, 1999.
[61] P.P. Mitra, S. Ogawa, X. Hu, and K. Ugurbil, "The nature of spatiotemporal changes in cerebral hemodynamics as manifested in functional magnetic resonance imaging," *Magn. Reson. Med.*, vol. 37, no. 4, pp. 511–518, 1997.
[62] T.H. Le and X.P. Hu, "Retrospective estimation and correction of physiological artifacts in fMRI by direct extraction of physiological activity from MR data," *Magn. Reson. Med.*, vol. 35, no. 3, pp. 290–298, 1996.
[63] J. Pfeuffer, P.F. Van de Moortele, K. Ugurbil, X.P. Hu, and G.H. Glover, "Correction of physiologically induced global off resonance effects in dynamic echo-planar and spiral functional imaging," *Magn. Reson. Med.*, vol. 47, no. 2, pp. 344–353, 2002.
[64] L. Zollei, L. Panych, E. Grimson, and W.M. Wells, "Exploratory identification of cardiac noise in fMRI images," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2003, Pt. 1*, vol. 2878, *Lecture Notes in Computer Science*, pp. 475–482, 2003.
[65] K.H. Chuang and J.H. Chen, "IMPACT: Image-based physiological artifacts estimation and correction technique for functional MRI," *Magn. Reson. Med.*, vol. 46, no. 2, pp. 344–353, 2001.
[66] X. Hu, T.H. Le, T. Parrish, and P. Erhard, "Retrospective estimation and correction of physiological fluctuation in functional MRI," *Magn. Reson. Med.*, vol. 34, no. 2, pp. 201–212, 1995.
[67] G.H. Glover, T.Q. Li, and D. Ress, "Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR," *Magn. Reson. Med.*, vol. 44, no. 1, pp. 162–167, 2000.
[68] K. Boesen, K. Rehm, K. Schaper, S. Stoltzner, R. Woods, E. Luders, and D. Rottenberg, "Quantitative comparison of four brain extraction algorithms," *NeuroImage*, vol. 22, no. 3, pp. 1255–1261, 2004.
[69] D.E. Rex, D.W. Shattuck, R.P. Woods, K.L. Narr, E. Luders, K. Rehm, S.E. Stolzner, D.A. Rottenberg, and A.W. Toga, "A meta-algorithm for brain extraction in MRI," *NeuroImage*, vol. 23, no. 2, pp. 625–637, 2004.
[70] F. Segonne, A.M. Dale, E. Busa, M. Glessner, D. Salat, H.K. Hahn, and B. Fischl, "A hybrid approach to the skull stripping problem in MRI," *NeuroImage*, vol. 22, no. 3, pp. 1060–1075, 2004.
[71] K. Rehm, K. Schaper, J. Anderson, R. Woods, S. Stoltzner, and D. Rottenberg, "Putting our heads together: A consensus approach to brain/non-brain segmentation in T1-weighted MR volumes," *NeuroImage*, vol. 22, no. 3, pp. 1262–1270, 2004.
[72] J. Ashburner and K.J. Friston, "Voxel-based morphometry—the methods," *NeuroImage*, vol. 11, no. 6, pt. 1, pp. 805–821, 2000.
[73] S.M. Smith, "Fast robust automated brain extraction," *Hum. Brain Map.*, vol. 17, no. 3, pp. 143–155, 2002.
[74] D.W. Shattuck, S.R. Sandor-Leahy, K.A. Schaper, D.A. Rottenberg, and R.M. Leahy, "Magnetic resonance image tissue classification using a partial volume model," *NeuroImage*, vol. 13, no. 5, pp. 856–876, 2001.
[75] B.D. Ward, "Intracranial Segmentation," Medical Coll. Wisconsin, Tech. Rep., 1999.
[76] A.M. Dale, B. Fischl, and M.I. Sereno, "Cortical surface-based analysis. I. Segmentation and surface reconstruction," *NeuroImage*, vol. 9, no. 2, pp. 179–194, 1999.
[77] H.K. Hahn, H.O. Peitgen, "The skull stripping problem in MRI solved by a single 3D watershed transform," presented at MICCAI 2000, Pittsburgh, PA, 2000.
[78] N. Kovacevic, N.J. Lobaugh, M.J. Bronskill, B. Levine, A. Feinstein, and S.E. Black, "A robust method for extraction and automatic segmentation of brain images," *NeuroImage*, vol. 17, no. 3, pp. 1087–1100, 2002.
[79] U. Vovk, F. Pernus, and B. Likar, "MRI intensity inhomogeneity correction by combining intensity and spatial information," *Phys. Med. Biol.*, vol. 49, pp. 4119–4133, 2004.
[80] C. Studholme, V. Cardenas, E. Song, F. Ezekiel, A. Maudsley, and M. Weiner, "Accurate template-based correction of brain MRI intensity distortion with application to dementia and aging," *IEEE Trans. Med. Imag.*, vol. 23, no. 1, pp. 99–110, 2004.
[81] E.B. Lewis and N.C. Fox, "Correction of differential intensity inhomogeneity in longitudinal MR images," *NeuroImage*, vol. 23, no. 1, pp. 75–83, 2004.
[82] J.B. Arnold, J.S. Liow, K.A. Schaper, J.J. Stern, J.G. Sled, D.W. Shattuck, A.J. Worth, M.S. Cohen, R.M. Leahy, J.C. Mazziotta, and D.A. Rottenberg, "Qualitative and quantitative evaluation of six algorithms for correcting intensity nonuniformity effects," *NeuroImage*, vol. 13, no. 5, pp. 931–943, 2001.
[83] J.G. Sled, A.P. Zijdenbos, and A.C. Evans, "A nonparametric method for automatic correction of intensity nonuniformity in MRI data," *IEEE Trans. Med. Imag.*, vol. 17, no. 1, pp. 87–97, 1998.
[84] S.M. Smith, M. Jenkinson, M.W. Woolrich, C.F. Beckmann, T.E. Behrens, H. Johansen-Berg, P.R. Bannister, M. De Luca, I. Drobnjak, D.E. Flitney, R.K. Niazy, J. Saunders, J. Vickers, Y. Zhang, N. De Stefano, J.M. Brady, and P.M. Matthews, "Advances in functional and structural MR image analysis and implementation as FSL," *NeuroImage*, vol. 23 Suppl. 1, pp. S208–S219, 2004.
[85] Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm," *IEEE Trans. Med. Imag.*, vol. 20, no. 1, pp. 45–57, 2001.
[86] M.R. Zaini, S.C. Strother, J.R. Anderson, J.S. Liow, U. Kjems, C. Tegeler, and S.G. Kim, "Comparison of matched BOLD and FAIR 4.0T-fMRI with [15O]water PET brain volumes," *Med. Phys.*, vol. 26, no. 8, pp. 1559–1567, 1999.
[87] B.A. Ardekani, A.H. Bachman, S.C. Strother, Y. Fujibayashi, and Y. Yonekura, "Impact of inter-subject image registration on group analysis of fMRI data," in *Proc. Int. Workshop Quantitation Biomedical Imaging PET MRI*, Osaka, Japan, 2004, vol. 1265.
[88] J. Talairach and P. Tournoux, *Co-Planar Stereotaxic Atlas of the Human Brain*. Stuttgart, Germany: Georg Theme Verlag, 1988.
[89] R.W. Cox, "AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages," *Comput. Biomed. Res.*, vol. 29, no. 3, pp. 162–173, 1996.

[90] R.P. Woods, S.T. Grafton, J.D. Watson, N.L. Sicotte, and J.C. Mazziotta, "Automated image registration: II. Intersubject validation of linear and nonlinear models," *J. Comput. Assisted Tomography*, vol. 22, no. 1, pp. 153–165, 1998.
[91] J. Ashburner and K.J. Friston, "Nonlinear spatial normalization using basis functions," *Hum. Brain Map.*, vol. 7, no. 4, pp. 254–266, 1999.
[92] U. Kjems, S.C. Strother, J. Anderson, I. Law, and L.K. Hansen, "Enhancing the multivariate signal of [15O] water PET studies with a new nonlinear neuroanatomical registration algorithm," *IEEE Trans. Med. Imag.*, vol. 18, no. 4, pp. 306–319, 1999.
[93] L. Collins and A. Evans, "Animal: Validation and applications of non-linear registration-based segmentation," *Int. J. Pattern Recognition Artificial Intell.*, vol. 8, no. 11, pp. 1271–1294, 1997.
[94] D.L. Collins, C.J. Holmes, T.M. Peters, and A.C. Evans, "Automatic 3–D model-based neuroanatomical segmentation.," *Hum. Brain Map.*, vol. 3, no. 3, pp. 190–208, 1995.
[95] G.E. Christensen, S.C. Joshi, and M.I. Miller, "Volumetric transformation of brain anatomy," *IEEE Trans. Med. Imag.*, vol. 16, no. 6, pp. 864–877, 1997.
[96] G.E. Christensen and H.J. Johnson, "Consistent image registration," *IEEE Trans. Med. Imag.*, vol. 20, no. 7, pp. 568–582, 2001.
[97] P. Hellier, C. Barillot, E. Memin, and P. Perez, "Hierarchical estimation of a dense deformation field for 3–D robust registration," *IEEE Trans. Med. Imag.*, vol. 20, no. 5, pp. 388–402, 2001.
[98] J.P. Thirion, "Image matching as a diffusion process: an analogy with Maxwell's demons," *Med. Image Anal.*, vol. 2, no. 3, pp. 243–260, 1998.
[99] P. Kochunov, J. Lancaster, P. Thompson, A. Boyer, J. Hardies, and P. Fox, "Evaluation of octree regional spatial normalization method for regional anatomical matching," *Hum. Brain Map.*, vol. 11, no. 3, pp. 193–206, 2000.
[100] M. Brett, I.S. Johnsrude, and A.M. Owen, "The problem of functional localization in the human brain," *Nat. Rev. Neurosci.*, vol. 3, no. 3, pp. 243–249, 2002.
[101] R.P. Woods, "Characterizing volume and surface deformations in an atlas framework: Theory, applications, and implementation," *NeuroImage*, vol. 18, no. 3, pp. 769–788, 2003.
[102] A. Guimond, J. Meunier, and J. Thirion, "Average brain models: A convergence study," *Comput. Vis. Imag. Understanding*, vol. 77, no. 2, pp. 192–210, 2000.
[103] P. Kochunov, J.L. Lancaster, P. Thompson, R. Woods, J. Mazziotta, J. Hardies, and P. Fox, "Regional spatial normalization: Toward an optimal target," *J. Comput. Assisted Tomography*, vol. 25, no. 5, pp. 805–816, 2001.
[104] N. Kovacevic, J. Chen, J.G. Sled, J. Henderson, and M. Henkelman, "Deformation based representation of groupwise average and variability," in *Proc. Medical Image Computing and Computer-Assisted Intervention—MICCAI*, pt. 1, Saint-Malo, France, 2004,pp. 615–622.
[105] P. Hellier, C. Barillot, I. Corouge, B. Gibaud, G. Le Goualher, D.L. Collins, A. Evans, G. Malandain, N. Ayache, G.E. Christensen, and H.J. Johnson, "Retrospective evaluation of intersubject brain registration," *IEEE Trans. Med. Imag.*, vol. 22, no. 9, pp. 1120–1130, 2003.
[106] H.J. Johnson and G.E. Christensen, "Consistent landmark and intensity-based image registration," *IEEE Trans. Med. Imag.*, vol. 21, no. 5, pp. 450–461, 2002.
[107] P. Hellier and C. Barillot, "Coupling dense and landmark-based approaches for nonrigid registration," *IEEE Trans. Med. Imag.*, vol. 22, no. 2, pp. 217–227, 2003.
[108] A. Leow, C.L. Yu, S.J. Lee, S.C. Huang, H. Protas, R. Nicolson, K.M. Hayashi, A.W. Toga, and P.M. Thompson, "Brain structural mapping using a novel hybrid implicit/explicit framework based on the level-set method," *NeuroImage*, vol. 24, no. 3, pp. 910–927, 2005.
[109] V.A. Magnotta, H.J. Bockholt, H.J. Johnson, G.E. Christensen, and N.C. Andreasen, "Subcortical, cerebellar, and magnetic resonance based consistent brain image registration," *NeuroImage*, vol. 19, no. 2, pt. 1, pp. 233–245, 2003.
[110] B. Fischl, M.I. Sereno, R.B. Tootell, and A.M. Dale, "High-resolution intersubject averaging and a coordinate system for the cortical surface," *Hum. Brain Map.*, vol. 8, no. 4, pp. 272–284, 1999.
[111] D.C. Van Essen, "Surface-based approaches to spatial localization and registration in primate cerebral cortex," *NeuroImage*, vol. 23, suppl. 1, pp. S97–S107, 2004.
[112] I.D. Dinov, M.S. Mega, P.M. Thompson, R.P. Woods, W.L. Sumners de, E.L. Sowell, and A.W. Toga, "Quantitative comparison and analysis of brain image registration using frequency-adaptive wavelet shrinkage," *IEEE Trans Inform. Technol. Biomed.,* vol. 6, no. 1, pp. 73–85, 2002.
[113] F. Crivello, T. Schormann, N. Tzourio-Mazoyer, P.E. Roland, K. Zilles, and B.M. Mazoyer, "Comparison of spatial normalization procedures and their impact on functional maps," *Hum. Brain Map.,* vol. 16, no. 4, pp. 228–250, 2002.
[114] T. Schormann and K. Zilles, "Three-dimensional linear and nonlinear transformations: An integration of light microscopical and MRI data," *Hum. Brain Map.*, vol. 6, no. 5-6, pp. 339–347, 1998.
[115] J.B. Poline and B.M. Mazoyer, "Enhanced detection in brain activation maps using a multifiltering approach," *J Cereb. Blood Flow Metab.,* vol. 14, no. 4, pp. 639–642, 1994.
[116] K.J. Friston, K.J. Worsley, R.S.J. Frackowiak, J.C. Mazziotta, and A.C. Evans, "Assessing the significance of focal activations using their spatial extent," *Hum. Brain Map.,* vol. 1, no. 3, pp. 210–220, 1993.
[117] K.J. Worsley, S. Marrett, P. Neelin, and A.C. Evans, "Searching scale space for activation in PET images," *Hum. Brain Mapp.*, vol. 4, no. 1, pp. 74–90, 1996.
[118] P. Skudlarski, R.T. Constable, and J.C. Gore, "ROC analysis of statistical methods used in functional MRI: individual subjects," *NeuroImage*, vol. 9, no. 3, pp. 311–329, 1999.
[119] T.B. Parrish, D.R. Gitelman, K.S. LaBar, and M.M. Mesulam, "Impact of signal-to-noise on functional MRI," *Magn. Reson. Med.,* vol. 44, no. 6, pp. 925–932, 2000.
[120] M.J. Lowe and J.A. Sorenson, "Spatially filtering functional magnetic resonance imaging data," *Magn. Reson. Med.*, vol. 37, no. 5, pp. 723–729, 1997.
[121] S. LaConte, J. Anderson, S. Muley, J. Ashe, S. Frutiger, K. Rehm, L.K. Hansen, E. Yacoub, X. Hu, D. Rottenberg, and S. Strother, "The evaluation of preprocessing choices in single-subject BOLD fMRI using NPAIRS performance metrics," *NeuroImage,* vol. 18, no. 1, pp. 10–27, 2003.
[122] R. Nandy and D. Cordes, "Improving the spatial specificity of canonical correlation analysis in fMRI," *Magn. Reson. Med.*, vol. 52, no. 4, pp. 947–952, 2004.
[123] A.F. Sole, S.C. Ngan, G. Sapiro, X. Hu, and A. Lopez, "Anisotropic 2-D and 3-D averaging of fMRI signals," *IEEE Trans. Med. Imag.,* vol. 20, no. 2, pp. 86–93, 2001.
[124] C. Long, E.N. Brown, D. Manoach, and V. Solo, "Spatiotemporal wavelet analysis for functional MRI," *NeuroImage,* vol. 23, no. 2, pp. 500–516, 2004.
[125] A.M. Wink and J.B. Roerdink, "Denoising functional MR images: A comparison of wavelet denoising and Gaussian smoothing," *IEEE Trans. Med. Imag.,* vol. 23, no. 3, pp. 374–387, 2004.
[126] J.S. Liow, K. Rehm, S.C. Strother, J.R. Anderson, N. Morch, L.K. Hansen, K.A. Schaper, and D.A. Rottenberg, "Comparison of voxel- and volume-of-interest-based analyses in FDG PET scans of HIV positive and healthy individuals," *J. Nucl. Med.,* vol. 41, no. 4, pp. 612–621, 2000.
[127] M.J. McKeown and C.A. Hanlon, "A post-processing/region of interest (ROI) method for discriminating patterns of activity in statistical maps of fMRI data," *J. Neurosci. Methods,* vol. 135, no. 1-2, pp. 137–147, 2004.
[128] J.R. Moeller and S.C. Strother, "A regional covariance approach to the analysis of functional patterns in positron emission tomographic data," *J. Cereb. Blood Flow Metab.,* vol. 11, no. 2, pp. A121–A135, 1991.
[129] J.L. Andersson, "How to estimate global activity independent of changes in local activity," *NeuroImage,* vol. 6, no. 4, pp. 237–244, 1997.
[130] G.K. Aguirre, E. Zarahn, and M. D'Esposito, "The inferential impact of global signal covariates in functional neuroimaging analyses," *NeuroImage,* vol. 8, no. 3, pp. 302–306, 1998.
[131] A.E. Desjardins, K.A. Kiehl, and P.F. Liddle, "Removal of confounding effects of global signal in functional MRI analyses," *NeuroImage,* vol. 13, no. 4, pp. 751–758, 2001.
[132] M. Gavrilescu, M.E. Shaw, G.W. Stuart, P. Eckersley, I.D. Svalbe, and G.F. Egan, "Simulation of the effects of global normalization procedures in functional MRI," *NeuroImage,* vol. 17, no. 2, pp. 532–542, 2002.
[133] E. Bullmore, J. Fadili, V. Maxim, L. Sendur, B. Whitcher, J. Suckling, M. Brammer, and M. Breakspear, "Wavelets and functional magnetic resonance imaging of the human brain," *NeuroImage,* vol. 23, suppl. 1, pp. S234–S249, 2004.
[134] M.W. Woolrich, B.D. Ripley, M. Brady, and S.M. Smith, "Temporal autocorrelation in univariate linear modeling of FMRI data," *NeuroImage,* vol. 14, no. 6, pp. 1370–86, 2001.
[135] K.J. Friston, O. Josephs, E. Zarahn, A.P. Holmes, S. Rouquette, and J. Poline, "To smooth or not to smooth? Bias and efficiency in fMRI time-series analysis," *NeuroImage,* vol. 12, no. 2, pp. 196–208, 2000.
[136] M. Bianciardi, A. Cerasa, F. Patria, and G.E. Hagberg, "Evaluation of mixed effects in event-related fMRI studies: Impact of first-level design and filtering," *NeuroImage*, vol. 22, no. 3, pp. 1351–1370, 2004.
[137] A.M. Smith, B.K. Lewis, U.E. Ruttimann, F.Q. Ye, T.M. Sinnwell, Y. Yang, J.H. Duyn, and J.A. Frank, "Investigation of low frequency drift in fMRI signal," *NeuroImage,* vol. 9, no. 5, pp. 526–533, 1999.
[138] R.G. Wise, K. Ide, M.J. Poulin, and I. Tracey, "Resting fluctuations in arterial carbon dioxide induce significant low frequency variations in BOLD signal," *NeuroImage,* vol. 21, no. 4, pp. 1652–1664, 2004.
[139] A.P. Holmes, O. Josephs, C. Büchel, and K.J. Friston, "Statistical modeling of low-frequency confounds in fMRI. Neuroimage 5 (pt. 2 of 4):S480," *NeuroImage,* vol. 5, pt. 2, p. S480, 1997.
[140] J. Zhang, S. Pulapura, J. Anderson, and S.C. Strother, "Exploring the Optimization of fMRI Processing Pipelines within the NPAIRS Framework," in *10th Annual Meeting Org. Hum. Brain Mapp.*, Budapest Hungary, 2004 [Online]. Available: http://www.meetingassistant.com/ ohbm/meeting_plan/ohbm_mtg_index_search.php
[141] J. Tanabe, D. Miller, J. Tregellas, R. Freedman, and F.G. Meyer, "Comparison of detrending methods for optimal fMRI preprocessing," *NeuroImage,* vol. 15, no. 4, pp. 902–907, 2002.
[142] J.L. Marchini, and B.D. Ripley, "A new statistical approach to detecting significant activation in functional MRI," *NeuroImage,* vol. 12, no. 4, pp. 366–380, 2000.
[143] K. Fissell, E. Tseytlin, D. Cunningham, K. Iyer, C.S. Carter, W. Schneider, and J.D. Cohen, "Fiswidgets: A graphical computing environment for neuroimaging analysis," *Neuroinformatics,* vol. 1, no. 1, pp. 111–125, 2003.
[144] S.C. Strother, "A developer's commentary on Fiswidgets," *Neuroinformatics,* vol. 1, no. 1, pp. 131–133, 2003.
[145] D.E. Rex, J.Q. Ma, and A.W. Toga, "The LONI Pipeline Processing Environment," *NeuroImage,* vol. 19, no. 3, pp. 1033–1048, 2003.
[146] S.C. Strother, J. Anderson, L.K. Hansen, U. Kjems, R. Kustra, J. Sidtis, S. Frutiger, S. Muley, S. LaConte, and D. Rottenberg, "The quantitative evaluation of functional neuroimaging experiments: The NPAIRS data analysis framework," *NeuroImage,* vol. 15, no.4, pp. 747–771, 2002.
[147] U. Kjems, L.K. Hansen, J. Anderson, S. Frutiger, S. Muley, J. Sidtis, D. Rottenberg, and S.C. Strother, "The quantitative evaluation of functional neuroimaging experiments: Mutual information learning curves," *NeuroImage,* vol. 15, no. 4, pp. 772–786, 2002.